

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:38:24 ; Search time 143 Seconds  
(without alignments)

448.896 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 964

Sequence: 1 VGAPVGGVARALAHGVRALE.....TVQTNCNSLYPGHLGSRMA 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1582122 seqs, 356623098 residues

Total number of hits satisfying chosen parameters: 1582122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	100.0	180	10	US-09-899-046-20
2	964	100.0	180	10	US-09-899-046-22
3	964	100.0	180	10	US-09-878-281-20
4	964	100.0	180	10	US-09-878-281-22
5	964	100.0	180	10	US-09-873-224-20
6	964	100.0	180	10	US-09-873-224-22
7	964	100.0	208	9	US-09-973-025-30
8	964	100.0	208	10	US-09-899-303-30
9	964	100.0	208	10	US-09-995-808-30
10	964	100.0	208	10	US-09-995-860-30
11	964	100.0	208	10	US-09-995-791-30
12	964	100.0	208	16	US-10-321-798-30
13	964	100.0	209	14	US-10-128-590-91

14 964 100.0 209 14 US-10-128-587A-91 Sequence 91, Appl  
15 960 99.6 193 15 US-10-651-165-210 Sequence 210, Appl  
16 953 98.9 187 15 US-10-651-165-209 Sequence 209, Appl  
17 951 98.7 180 10 US-09-899-046-24 Sequence 24, Appl  
18 951 98.7 180 10 US-09-899-046-26 Sequence 26, Appl  
19 951 98.7 180 10 US-09-899-046-28 Sequence 28, Appl  
20 951 98.7 180 10 US-09-878-281-24 Sequence 24, Appl  
21 951 98.7 180 10 US-09-878-281-26 Sequence 26, Appl  
22 951 98.7 180 10 US-09-878-281-28 Sequence 28, Appl  
23 951 98.7 180 10 US-09-873-224-24 Sequence 24, Appl  
24 951 98.7 180 10 US-09-873-224-26 Sequence 26, Appl  
25 951 98.7 180 10 US-09-873-224-28 Sequence 28, Appl  
26 951 98.7 180 10 US-10-651-165-208 Sequence 208, Appl  
27 948 98.3 180 10 US-09-899-046-14 Sequence 14, Appl  
28 948 98.3 180 10 US-09-899-046-18 Sequence 18, Appl  
29 948 98.3 180 10 US-09-878-281-14 Sequence 14, Appl  
30 948 98.3 180 10 US-09-878-281-18 Sequence 18, Appl  
31 948 98.3 180 10 US-09-873-224-14 Sequence 14, Appl  
32 948 98.3 180 10 US-09-873-224-18 Sequence 18, Appl  
33 948 98.3 187 15 US-10-651-165-207 Sequence 207, Appl  
34 941 97.6 180 10 US-09-899-046-16 Sequence 16, Appl  
35 941 97.6 180 10 US-09-878-281-16 Sequence 16, Appl  
36 941 97.6 180 10 US-09-873-224-16 Sequence 16, Appl  
37 791 82.1 319 15 US-10-651-165-211 Sequence 211, Appl  
38 753 78.1 193 10 US-09-899-046-176 Sequence 176, Appl  
39 753 78.1 193 10 US-09-878-281-176 Sequence 176, Appl  
40 733 78.1 193 10 US-09-873-224-176 Sequence 176, Appl  
41 753 78.1 193 15 US-10-651-165-221 Sequence 221, Appl  
42 744 77.2 191 10 US-09-899-046-121 Sequence 121, Appl  
43 744 77.2 191 10 US-09-878-281-121 Sequence 121, Appl  
44 744 77.2 191 10 US-09-873-224-121 Sequence 121, Appl  
45 734 76.1 450 15 US-10-651-165-189 Sequence 189, Appl

#### ALIGNMENTS

RESULT 1  
US-09-899-046-20  
; Sequence 20, Application US/09899046  
; Publication No. US20030008274A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; genotype for diagnosis, prophylaxis and therapy.  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION NUMBER: US/09/899,046  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-899-046-20

Query Match 100.0%; Score 964; DB 10; Length 180;

Best Local Similarity 100.0%; Pred. No. 1.7e-96; Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60

Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60

QY 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
DB 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 2

US-09-899-046-22  
; Sequence 22, Application US/09899046  
; Publication No. US20030008274A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/899,046  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-899-046-22

Query Match 100.0%; Score 964; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.7e-96;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
QY 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
DB 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 3

US-09-898-281-20  
; Sequence 20, Application US/09878281  
; Publication No. US20030032005A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/878,281  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455

; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-878-281-20

Query Match 100.0%; Score 964; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.7e-96;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
QY 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
DB 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 4

US-09-878-281-22  
; Sequence 22, Application US/09878281  
; Publication No. US20030032005A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/878,281  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-878-281-22

Query Match 100.0%; Score 964; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.7e-96;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
QY 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
DB 61 LYVLNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

RESULT 5  
US-09-873-224-20

```
; Sequence 20, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
;               genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics sa.
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-873-224-22

Query Match      100.0%; Score 964; DB 10; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.7e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
   |||||||
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
   |||||||

QY 61 LVLTNDCSNLSIVVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
   |||||||
Db 61 LVLTNDCSNLSIVVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
   |||||||

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
   |||||||
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
   |||||||

RESULT 7
US-09-873-025-30
; Sequence 30, Application US/09973025
; Publication No. US20020182706A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
;               BOSMAN, FONS
;               DE MARTYNOFF, GUY
;               BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
;               PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/973,025
; FILING DATE: 10-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205

; Sequence 20, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
;               genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



```
; ORGANISM: Hepatitis C virus
US-09-995-860-30

Query Match      100.0%; Score 964; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Db 22 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81
QY 61 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
Db 82 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 141
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 11
US-09-995-791-30
; Sequence 30, Application US/09995791
; Publication No. US20030147918A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-68
; CURRENT APPLICATION NUMBER: US/09/995,791
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 30
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-995-791-30

Query Match      100.0%; Score 964; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Db 22 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81
QY 61 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
Db 82 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 141
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 12
US-10-128-798-30
; Sequence 30, Application US/10321798
; Publication No. US20040126395A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-93
; CURRENT APPLICATION NUMBER: US/10/321,798
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 60/418,358
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 10/020,510
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn 3.1
```

```
; SEQ ID NO 30
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-321-798-30

Query Match      100.0%; Score 964; DB 16; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Db 22 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81
QY 61 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
Db 82 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 141
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 13
US-10-128-590-91
; Sequence 91, Application US/10128590
; Publication No. US20030108561A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Core-glycosylated HCV envelope proteins
; FILE REFERENCE: 135 PCT
; CURRENT APPLICATION NUMBER: US/10/128,590
; CURRENT FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
; LENGTH: 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-128-590-91

Query Match      100.0%; Score 964; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Db 23 VGAPVGGVARALAHGVRALGDALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 82
QY 61 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
Db 83 LYVLTNDCSNSISIVYEADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 142
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 143 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 202

RESULT 14
US-10-128-587A-91
; Sequence 91, Application US/10128587A
; Publication No. US20030152940A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Constructs and methods for expression of recombinant HCV envelope
; FILE REFERENCE: 134 PCT
; CURRENT APPLICATION NUMBER: US/10/128,587A
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
; LENGTH: 209
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: hepatitis C virus
US-10-128-587A-91

Query Match      100.0%; Score 964; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60
Db 23 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 82
QY 61 LYVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTPTVAVKYVGATTASIR 120
Db 83 LYVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTPTVAVKYVGATTASIR 142
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHOTVOTCNCSLYPGHLGHRMA 180
Db 143 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHOTVOTCNCSLYPGHLGHRMA 202
```

```
RESULT 15
US-10-651-165-210
; Sequence 210, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
; LENGTH: 193
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-210
```

```
Query Match      99.6%; Score 960; DB 15; Length 193;
Best Local Similarity 98.9%; Pred. No. 5.1e-96;
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60
Db 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 73
QY 61 LYVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTPTVAVKYVGATTASIR 120
Db 74 LYVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTPTVAVKYVGATTASIR 133
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHOTVOTCNCSLYPGHLGHRMA 180
Db 134 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHOTVOTCNCSLYPGHLGHRMA 193
```

Search completed: December 1, 2004, 23:48:10  
Job time : 144 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:35:49 ; Search time 39 Seconds  
(without alignments)  
306.083 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 964

Sequence: 1 VGAPVGGVARALAHGVRALE.....TVQTCNCSLYPGHLSGHRMA 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PGTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	100.0	180	4	US-09-878-281A-20
2	964	100.0	180	4	US-09-878-281A-22
3	964	100.0	208	3	US-08-612-973-30
4	964	100.0	208	3	US-08-927-597-30
5	960	99.6	193	4	US-08-635-886C-210
6	960	99.6	193	4	US-08-974-690C-210
7	953	98.9	187	4	US-08-635-886C-209
8	953	98.9	187	4	US-08-974-690C-209
9	951	98.7	180	4	US-09-878-281A-24
10	951	98.7	180	4	US-09-878-281A-26
11	951	98.7	180	4	US-09-878-281A-28
12	951	98.7	187	4	US-08-635-886C-208
13	951	98.7	187	4	US-08-974-690C-208
14	948	98.3	180	4	US-09-878-281A-14
15	948	98.3	180	4	US-09-878-281A-18
16	948	98.3	187	4	US-08-635-886C-207
17	948	98.3	187	4	US-08-974-690C-207
18	941	97.6	180	4	US-09-878-281A-16
19	791	82.1	319	4	US-08-635-886C-211
20	791	82.1	319	4	US-08-974-690C-211
21	753	78.1	193	4	US-08-635-886C-221
22	753	78.1	193	4	US-08-974-690C-221
23	753	78.1	193	4	US-09-878-281A-176
24	744	77.2	191	4	US-09-878-281A-121
25	739	76.7	3010	3	US-09-014-416-3
26	739	76.7	3011	3	US-09-014-416-5
27	734	76.1	450	4	US-08-635-886C-189

#### ALIGNMENTS

##### RESULT 1

US-09-878-281A-20  
; Sequence 20, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Innogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy  
; FILE OF INVENTION: 35  
; CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 20  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-09-878-281A-20

Query Match 100.0%; Score 964; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 8e+104;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VGAPVGGVARALAHGVRALE	GINPATGNLPGCSFIFLLALFSLIHPAASLEWRNTSG	60
Db	1	VGAPVGGVARALAHGVRALE	GINPATGNLPGCSFIFLLALFSLIHPAASLEWRNTSG	60
QY	61	LYVLNDGNSISIVYEAD	VILHTPGCIPCVQDGNSTCMTPTPTVAVKVGGATTASIR	120
Db	61	LYVLNDGNSISIVYEAD	VILHTPGCIPCVQDGNSTCMTPTPTVAVKVGGATTASIR	120
QY	121	SHVDLLVGAATMCSALYV	GDGCGAVFLNGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA	180
Db	121	SHVDLLVGAATMCSALYV	GDGCGAVFLNGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA	180

##### RESULT 2

US-09-878-281A-22  
; Sequence 22, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Innogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy  
; FILE OF INVENTION: 35  
; CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 22

```
; LENGTH: 180
; TYPE: PROT
; ORGANISM: hepatitis C virus
US-09-878-281A-22

Query Match      100.0%; Score 964; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 8e-104;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
QY 61 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
DB 61 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
QY 121 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 121 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 121 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

RESULT 3
US-08-612-973-30
; Sequence 30, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-612-973-30

Query Match      100.0%; Score 964; DB 3; Length 208;
Best Local Similarity 100.0%; Pred. No. 9.9e-104;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 81
QY 61 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
DB 82 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 141
QY 121 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 142 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 5
US-08-927-597-30
; Sequence 30, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-927-597-30

Query Match      100.0%; Score 964; DB 3; Length 208;
Best Local Similarity 100.0%; Pred. No. 9.9e-104;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 81
QY 61 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 120
DB 82 LYVLTDNCSSNIIYVYADDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR 141
QY 121 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 142 SHVDLLVGAATMCSALYVGDGMCAGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 201
```



```
US-08-635-886C-210
; Sequence 210, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; PRIOR FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-210

Query Match 99.6%; Score 960; DB 4; Length 193;
Best Local Similarity 98.9%; Pred. No. 2.6e-103;
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 73

QY 61 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 120
DB 74 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 133

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 134 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 193

RESULT 6
US-08-974-690C-210
; Sequence 210, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-210

Query Match 99.6%; Score 960; DB 4; Length 193;
Best Local Similarity 98.9%; Pred. No. 2.6e-103;
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 73

QY 61 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 120
DB 74 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 133

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 134 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 193

RESULT 7
US-08-635-886C-209
; Sequence 209, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; PRIOR FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-209

Query Match 98.9%; Score 953; DB 4; Length 187;
Best Local Similarity 98.9%; Pred. No. 1.6e-102;
Matches 178; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 8 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 67

QY 61 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 120
DB 68 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 127

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 128 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 187

RESULT 8
US-08-974-690C-209
; Sequence 209, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-209

Query Match 99.6%; Score 960; DB 4; Length 193;
Best Local Similarity 98.9%; Pred. No. 2.6e-103;
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 73

QY 61 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 120
DB 74 LVLTNDCSNLSIVYEADDVILHTPGCIPCVQDGNSTCTWTPTVAVKYVGATTASIR 133

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 134 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 193
```

; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-C8-974-690C-209

Query Match	98.9%	Score 953;	DB 4;	Length 187;
Best Local Similarity	98.9%	Pred. No. 1.6e-102;		
Matches 178; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIIHPAASLEWRNTSG	60	
DB	8	VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIIHPAASLEWRNTSG	67	
QY	61	LYVLNDNCNSSIYYEADDVILHTPGICPVQDGNSTSTCTWPTVTVAKVVGATTASIR	120	
DB	68	LYVLNDNCNSSIYYEADDVILHTPGVPCVQDGNSTSTCTWPTVTVAKVVGATTASIR	127	
QY	121	SHVDLLVGAATMCSSALYYGDMCAVFLVGOAFTFRPRRHQTQTCNCSLYPGH-SGHRMA	180	
DB	128	SHVDLLVGAATMCSSALYYGDMCAVFLVGOAFTFRPRRHQTQTCNCSLYPGH-SGHRMA	187	

RESULT 9  
US-09-878-281A-24  
; Sequence 24, Application US/09878281A  
; Patent No. 6762024

```

/ Patent NO.: 0762024
/ GENERAL INFORMATION:
/ APPLICANT: Innogenetics N.V.
/ TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
/ TITLE OF INVENTION: and therapy
/ FILE REFERENCE: 35
/ CURRENT APPLICATION NUMBER: US/09/878,281A
/ CURRENT FILING DATE: 2001-06-12
/ NUMBER OF SEQ ID NOS: 284
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 24
/ LENGTH: 180
/ TYPE: PRT
/ ORGANISM: hepatitis C virus
/ US-09-878-281A-24

```

```

RESULT 10
US/09-878-281A-26
; Sequence 26, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepat
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 180
; TYPE: PRT

```

itis C virus genotypes for diagnosis, proph

; ORGANISM: hepatitis C virus  
US-09-878-281A-26

Query Match	98.7%;	Score 951;	DB 4;	Length 180;
Best Local Similarity	97.8%;	Pred. No. 2.6e-102;		
Matches 176;	Conservative 2;	Mismatches 102;	Indels 0;	Gaps 0
Qy	1	VCAPGGVARALAGVRALEDGINFATGNLPGCSFSIFLLALFSCULHPAALEWRNTSG	60	
Db	1	VCAPGGVARALAGVRALEDGINFATGNLPGCSFSIFLLALFSCULHPAALEWRNTSG	60	
Qy	61	LYVLITDCSNSIIYVEADDVILHTPGICPCVQDGNSTCWTPTPTVAVKYVGATTASIR	120	
Db	61	LYVLITDCSNSIIYVEADDVILHAPGCVPCVQDGNSTCWTPTPTVAVKYVGATTASIR	120	
Qy	121	SHVDLLVGAATMCALYVDMCGAVFLVGOAFTFRPRRHTQVOTNCSLYPGHLSGRMA	180	
Db	121	SHVDLLVGAATMCALYVDMCGAVFLVGOAFTFRPRRHTQVOTNCSLYPGHLSGRMA	180	

RESULT 11  
US-09-878-281A-28  
; Sequence 28, Application US/09878281A  
; Patent No. 6762024

```

/ Patent NO.: 6762024
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Innogenetics N.V.
/
/ TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
/
/ TITLE OF INVENTION: and therapy
/
/ FILE REFERENCE: 35
/
/ CURRENT APPLICATION NUMBER: US/09/878,281A
/
/ CURRENT FILING DATE: 2001-06-12
/
/ NUMBER OF SEQ ID NOS: 284
/
/ SOFTWARE: PatentIn version 3.1
/
/ SEQ ID NO 28
/
/ LENGTH: 180
/
/ TYPE: PRT
/
/ ORGANISM: hepatitis C virus
/
/ US-09-878-281A-28
/

```

RESULT 12  
US-08-635-886C-208  
: Sequence 208 Application IIS/08635886C

```

/ Sequence 208, Application US/08635886C
/ Patent No. 6555114
/
/ GENERAL INFORMATION:
/
/ APPLICANT: LEROUX-ROELS, Geert
/ APPLICANT: DELEYS, Robert
/ APPLICANT: MAERTENS, Geert
/
/ TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
/
/ TITLE OF INVENTION: VIRUS
/
/ FILE REFERENCE: 2752-18
/
/ CURRENT APPLICATION NUMBER: US/08/635.886C
/
/ CURRENT FILING DATE: 1996-04-25
/
/ PRIOR APPLICATION NUMBER: PCT/EP94/03555
/
/ PRIOR FILING DATE: 1994-10-28
/
/ PRIOR APPLICATION NUMBER: EP 93402718.6
/
/ PRIOR FILING DATE: 1993-11-04
/

```

```
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 208
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-08-635-886C-208

Query Match      98.7%; Score 951; DB 4; Length 187;
Best Local Similarity 97.8%; Pred. No. 2.7e-102;
Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Db 8 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAAGLEWRNTSG 67
Qy 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Db 68 LVYLTNDCNSSIIVYEADDVILHAFGCVPCVDGNSTCWTPTVAVKYVGATTASIR 127
Qy 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180
Db 128 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 187

RESULT 13
US-08-974-690C-208
; Sequence 208, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/E994/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 208
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-08-974-690C-208

Query Match      98.7%; Score 951; DB 4; Length 187;
Best Local Similarity 97.8%; Pred. No. 2.7e-102;
Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Db 8 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAAGLEWRNTSG 67
Qy 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Db 68 LVYLTNDCNSSIIVYEADDVILHAFGCVPCVDGNSTCWTPTVAVKYVGATTASIR 127
Qy 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180
Db 128 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 187

RESULT 14
US-09-878-281A-14
; Sequence 14, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
```

```
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-09-878-281A-14

Query Match      98.3%; Score 948; DB 4; Length 180;
Best Local Similarity 97.2%; Pred. No. 5.8e-102;
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Qy 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Db 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Qy 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180
Db 121 RHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180

RESULT 15
US-09-878-281A-18
; Sequence 18, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-09-878-281A-18

Query Match      98.3%; Score 948; DB 4; Length 180;
Best Local Similarity 97.2%; Pred. No. 5.8e-102;
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Qy 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Db 61 LVYLTNDCNSSIIVYEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Qy 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180
Db 121 RHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180

Search completed: December 1, 2004, 23:45:42
Job time : 40 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:26:38 ; Search time 152 Seconds  
(without alignments)

424.811 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 964

Sequence: 1 VGAPVGVARALAHGVRALE.....TVQTCNSLYPGHLSGRMA 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	100.0	180	2	AAR63280 Polypepti
2	964	100.0	180	2	AAR63281 Polypepti
3	964	100.0	208	5	AAO18670 Hepatitis
4	964	100.0	208	7	ADD55538 Hepatitis
5	964	100.0	208	8	ADP71120 HCV E1 pr
6	964	100.0	209	6	ADP55561 Hepatitis
7	964	100.0	209	6	AAE32874 Hepatitis
8	951	98.7	180	2	AAR63282 Polypepti
9	951	98.7	180	2	AAR63284 Polypepti
10	951	98.7	180	2	AAR63283 Polypepti
11	948	98.3	180	2	AAR63279 Polypepti
12	948	98.3	180	2	AAR63277 Polypepti
13	941	97.6	180	2	AAR63278 Polypepti
14	784.5	81.4	3023	2	AAR94462 Hepatitis
15	753	78.1	193	2	AAR63365 Hepatitis
16	744	77.2	191	2	AAR63332 Hepatitis
17	741	76.9	603	2	AAR12600 PT-NANBH
18	739	76.7	3010	2	AAR98022 Infectiou
19	739	76.7	3010	4	AAB31170 Amino aci
20	739	76.7	3010	8	ADO36227 Hepatitis
21	739	76.7	3010	8	ADO79401 Hepatitis
22	739	76.7	3011	2	AAR98021 Infectiou
23	737	76.5	1031	2	AAR54067 Non-A, no
24	737	76.5	1031	2	AAR98362 5'UTR/COR
25	733	76.0	326	2	AAR68461 Protein e

#### ALIGNMENTS

##### RESULT 1

AAR63280  
ID AAR63280 standard; protein; 180 AA.

XX AAR63280;  
AC AC  
XX 25-MAR-2003 (revised)  
DT 01-AUG-1995 (first entry)  
DT  
XX  
DE Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
XX  
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
KW classification; immunisation; prophylaxis; serotyping.  
XX  
OS Hepatitis C virus type 3a.

XX

PN WO9425601-A2.

XX

PD 10-NOV-1994.

XX

PF 27-APR-1994; 94WO-EP001323.

XX

PR 27-APR-1993; 93EP-00401099.

XX

PR 05-AUG-1993; 93EP-00402019.

XX

PI (INNO-) INNOGENETICS NV SA.

XX

PI Maertens G, Stuyver L;

XX

XX WPI; 1994-358277/44.

DR N-PSDB; AAQ78032.

XX

XX New polynucleotide sequences from hepatitis C virus - and related

PT vectors, polypeptide(s) and antibodies, useful for immunisation,

PT treatment, diagnosis and typing of HCV isolates.

XX

XX Claim 11; Page 110-111; 404pp; English.

XX

XX Compositions comprising at least 5, and pref. 8 or more contiguous

CC nucleotides selected from an HCV type 3 genomic sequence, more

26 733 76.0 423 2 AAW68465 Protein e  
27 733 76.0 513 2 AAR24087 NANB hepa  
28 733 76.0 729 2 AAM47264 Cuticle p  
29 733 76.0 733 2 AAR38279 NANB hepa  
30 733 76.0 845 2 AAW68466 Protein e  
31 733 76.0 3010 4 AAB59174 Protein e  
32 731 75.8 193 2 AAR63368 Hepatitis  
33 731 75.8 246 2 AAR20723 C10-E15 N  
34 731 75.8 310 2 AAR20720 C10-E12 N  
35 731 75.8 470 2 AAR34478 Encoded b  
36 731 75.8 470 2 AAR34477 Encoded b  
37 731 75.8 470 2 AAR34476 Encoded b  
38 731 75.8 2984 4 AAE00449 Hepatitis  
39 731 75.8 3011 2 AAW77398 Hepatitis  
40 731 75.8 3011 5 AAO99290 Hepatitis  
41 731 75.8 3011 6 AAO26784 Protein d  
42 731 75.8 3011 6 ABU61849 HCV-H. 8/  
43 730 75.7 321 2 AAR25115 Non-A, No  
44 729 75.6 967 2 AAR79222 pHCV141-e  
45 729 75.6 1006 2 AAW12715 HCV genom

CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated BR36-9-13.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 180 AA;

Query Match 100.0%; Score 964; DB 2; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-95;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 QY 61 LYVLINDCSNLSIVYEADDDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 DB 61 LYVLINDCSNLSIVYEADDDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGDWCGAVFLVGOAFTPRRRHQTVCNCSLYPGHLSGHRMA 180  
 DB 121 SHVDLLVGAATMCSALYVGDWCGAVFLVGOAFTPRRRHQTVCNCSLYPGHLSGHRMA 180

RESULT 2  
 AAR63281  
 ID AAR63281 standard; protein; 180 AA.  
 XX  
 AC AAR63281;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 01-AUG-1995 (first entry)  
 XX  
 DE Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
 XX  
 KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
 KW classification; immunisation; prophylaxis; serotyping.  
 XX  
 OS Hepatitis C virus type 3a.  
 XX  
 PN WO9425601-A2.  
 XX  
 PD 10-NOV-1994.  
 XX  
 PF 27-APR-1994; 94WO-EP001323.  
 XX  
 PR 27-APR-1993; 93EP-00401099.  
 PR 05-AUG-1993; 93EP-00402019.  
 XX  
 PA (INNO-) INNOGENETICS NV SA.  
 XX  
 PI Maertens G, Stuyver L;  
 XX  
 DR WPI; 1994-358277/44.  
 DR N-PSDB; AAQ78033.  
 XX

New polynucleotide sequences from hepatitis C virus - and related  
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
 PT treatment, diagnosis and typing of HCV isolates.  
 XX

Claim 11; Page 112-113; 404pp; English.  
 XX  
 CC Compositions comprising at least 5, and pref. 8 or more contiguous  
 CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-957 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
 CC positions 8023-8235 of the NS5 region of the B36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated BR36-9-20.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 180 AA;

Query Match 100.0%; Score 964; DB 2; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-95;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 QY 61 LYVLINDCSNLSIVYEADDDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 DB 61 LYVLINDCSNLSIVYEADDDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGDWCGAVFLVGOAFTPRRRHQTVCNCSLYPGHLSGHRMA 180  
 DB 121 SHVDLLVGAATMCSALYVGDWCGAVFLVGOAFTPRRRHQTVCNCSLYPGHLSGHRMA 180

RESULT 3  
 AAO18670  
 ID AAO18670 standard; protein; 208 AA.  
 XX  
 AC AAO18670;  
 XX  
 DT 24-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus clone HCCI62 E1 protein.  
 XX  
 KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; virucide;  
 KW immunostimulant; vaccine.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO200255548-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 11-JAN-2002; 2002WO-EP000219.  
 XX  
 PR 11-JAN-2001; 2001US-0260669P.  
 PR 30-AUG-2001; 2001US-0315768P.  
 XX  
 PA (INNO-) INNOGENETICS NV.  
 XX  
 PI Maertens G, Bosman F, Buysse M;  
 XX  
 DR WPI; 2002-599657/64.  
 DR N-PSDB; AAL48929.  
 XX

New therapeutic vaccine compositions comprising at least one purified  
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric  
 PT recombinant envelope protein E1 or E2, useful for immunizing humans from  
 PT HCV infection.  
 XX

Disclosure; Page 183; 243pp; English.  
 XX  
 CC The present invention relates to new therapeutic vaccine compositions for  
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a  
 CC composition containing at least one purified recombinant HCV single or  
 CC specific oligomeric recombinant envelope proteins selected from an E1 and  
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are  
 CC useful for inducing HCV-specific antibodies or for immunising humans  
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as  
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody  
 CC tests, for raising antibodies, in the preparation of medicament, and for  
 CC in vitro monitoring of HCV disease or prognosing the response to

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
|||  
Dh 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 81

CC optionally, a pharmaceutical adjuvant. The HCV vaccine composition is  
 CC useful for reducing liver disease (such as liver fibrosis or its  
 CC progression), serum ALT levels, steatosis, or anti-E2 immunoreactivity in  
 CC the liver in a chronic HCV-infected mammal, or for treating a chronic HCV  
 CC -infected mammal, particularly human. The HCV E1 proteins are useful for  
 CC in vitro monitoring HCV disease or prognosing the response to treatment  
 CC of patients suffering from HCV infection. The present sequence is an HCV  
 CC E1 protein (or fragment).

XX  
 SQ Sequence 208 AA;

Query Match 100.0%; Score 964; DB 8; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-95;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81  
 QY 61 LYVLTNDCSNSIVYEADVVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 DB 82 LYVLTNDCSNSIVYEADVVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 141  
 QY 121 SHVDLLVGAATWCSALYVGDGCGAVFLVQGFATPRRHHQVOTCNCSLYPGHLGHRMA 180  
 DB 142 SHVDLLVGAATWCSALYVGDGCGAVFLVQGFATPRRHHQVOTCNCSLYPGHLGHRMA 201

RESULT 6  
 ID ABP55561 standard; protein; 209 AA.  
 XX  
 AC ABP55561;

DT 19-FEB-2003 (first entry)

DE Hepatitis C virus protein SEQ ID NO:91.

XX Hepatitis C virus; HCV; envelope protein; hepatotropic; virucide;  
 KW antiinflammatory; gene therapy; vaccine.

XX Hepatitis C virus.

XX WO200285932-A2.

XX 31-OCT-2002.

XX 24-APR-2002; 2002WO-BE000062.

XX 24-APR-2001; 2001EP-00870088.

XX 17-JUL-2001; 2001US-0305604P.

XX (INNO-) INNOGENETICS NV.

XX Sablon E, Van Broekhoven A, Bosman A, Depla E, Deschamps G;

XX WPI; 2003-093095/08.

XX New recombinant nucleic acids for expressing Hepatitis C virus (HCV)  
 PT envelope proteins in eukaryotic cells, comprising a sequence encoding a  
 PT protein having an avian lysozyme leader peptide joined to the HCV  
 PT envelope protein.

XX Claim 5; Page 310-311; 319pp; English.

XX The present invention describes a recombinant nucleic acid (I) comprising  
 CC a nucleotide sequence encoding a protein having an avian lysozyme leader  
 CC peptide, or its functional equivalent, joined to a Hepatitis C virus  
 CC (HCV) envelope protein or its part. Also described: (1) a vector  
 CC comprising the recombinant nucleic acid; (2) a host cell comprising the  
 CC recombinant nucleic acid or the vector; and (3) a method for producing  
 CC HCV envelope protein or its part in a host cell, comprising transforming  
 CC the host cell with the recombinant nucleic acid or with the vector, where

CC the host cell is capable of expressing the protein cited above. (I) has  
 CC hepatotropic, virucide and antiinflammatory activities, and can be used  
 CC in gene therapy and vaccines. The recombinant nucleic acid is useful for  
 CC efficient expression of Hepatitis C virus envelope proteins in eukaryotic  
 CC cells, such as yeast cells. The HCV envelope proteins may be used as a  
 CC vaccine, or for incorporation into an immunoassay for the detection of  
 CC anti-HCV antibodies, and/or genotyping of HCV, for prognosing or  
 CC monitoring of HCV disease, or as a therapeutic agent. The method is used  
 CC for producing HCV envelope protein or its part in a host cell. ABQ84197  
 CC to ABQ84253 and ABP55528 to ABP55569 represent sequences used in the  
 CC exemplification of the present invention

XX  
 SQ Sequence 209 AA;

Query Match 100.0%; Score 964; DB 6; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-95;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 23 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 82  
 QY 61 LYVLTNDCSNSIVYEADVVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
 DB 83 LYVLTNDCSNSIVYEADVVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 142  
 QY 121 SHVDLLVGAATWCSALYVGDGCGAVFLVQGFATPRRHHQVOTCNCSLYPGHLGHRMA 180  
 DB 143 SHVDLLVGAATWCSALYVGDGCGAVFLVQGFATPRRHHQVOTCNCSLYPGHLGHRMA 202

RESULT 7  
 AAE32874  
 ID AAE32874 standard; protein; 209 AA.  
 XX  
 AC AAE32874;

DT 24-MAR-2003 (first entry)

XX Hepatitis C virus protein #9.

XX Hepatitis C virus; HCV; envelope protein; virucide; immunostimulant;  
 KW immune response; T-cell; therapy; infection; pharmaceutical; vaccine.

XX Hepatitis C virus.

XX WO200286101-A2.

XX 31-OCT-2002.

XX 24-APR-2002; 2002WO-BE000064.

XX 24-APR-2001; 2001EP-00870088.

XX 17-JUL-2001; 2001US-0305604P.

XX (INNO-) INNOGENETICS NV.

XX Depla E, Bosman A, Deschamps G, Sablon E, Suckow M, Samson I;

XX Verheyden G;

XX WPI; 2003-103409/09.

XX New Hepatitis C virus (HCV) envelope protein with N-glycosylation  
 PT site(s), useful as a vaccine for inducing a HCV-specific immune response  
 PT or HCV-specific antibodies, particularly for preventing or treating HCV  
 PT infection.

XX Claim 10; Page 343-344; 355pp; English.

XX The invention relates to Hepatitis C virus (HCV) envelope protein with N-  
 CC glycosylation site(s). The HCV envelope protein, or the pharmaceutical  
 CC composition comprising the envelope protein, is useful as a medicament or  
 CC a vaccine, particularly for inducing a HCV-specific immune response,



CC inducing HCV-specific antibodies or inducing a T-cell function in a  
CC mammal. The protein is particularly useful for preventing, treating or  
CC diagnosing HCV infection. It is also useful for detecting the presence of  
CC anti-HCV antibodies in a sample. The present sequence is Hepatitis C  
CC virus protein used in the invention

XX  
SQ Sequence 209 AA;

Query Match 100.0%; Score 964; DB 6; Length 209;  
Best Local Similarity 100.0%; Pred. No. 5.7e-95;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINPATNLPCCSFIFLLALFSCLIHPAASLEWRNTSG 60  
DB 23 VGAPVGGVARALAHGVRALEDGINPATNLPCCSFIFLLALFSCLIHPAASLEWRNTSG 82

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120  
DB 83 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 142

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQVQTCNSLYPGHLSGRMA 180  
DB 143 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQVQTCNSLYPGHLSGRMA 202

## RESULT 8

AAR63282  
ID AAR63282 standard; protein; 180 AA.

XX  
AC AAR63282;

XX 25-MAR-2003 (revised)

DT 01-AUG-1995 (first entry)

XX Polypeptide encoded by hepatitis C virus Core/E1 sequence.

DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
XX  
KW classification; immunisation; prophylaxis; serotyping.

XX  
OS Hepatitis C virus type 3a.

XX  
PN WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

DR N-PSDB; AAQ78034.

XX New polynucleotide sequences from hepatitis C virus - and related  
PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
PT treatment, diagnosis and typing of HCV isolates.

PS - Claim 11; Page 114-115; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
CC nucleotides selected from an HCV type 3 genomic sequence, more  
CC particularly (i) the region spanning positions 417-957 of the Core/E1  
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
CC for immunisation against HCV, for the detection of antibodies directed  
CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
CC region of HCV subtype 3a and is taken from a clone designated BR33-1-10.  
CC (Updated on 25-MAR-2003 to correct PN field.)

XX  
SQ Sequence 180 AA;

Query Match 98.7%; Score 951; DB 2; Length 180;  
Best Local Similarity 97.8%; Pred. No. 1.2e-93;  
Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINPATNLPCCSFIFLLALFSCLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALEDGINPATNLPCCSFIFLLALFSCLIHPAASLEWRNTSG 60

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120  
DB 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQVQTCNSLYPGHLSGRMA 180  
DB 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQVQTCNSLYPGHLSGRMA 180

## RESULT 9

AAR63284  
ID AAR63284 standard; protein; 180 AA.

XX  
AC AAR63284;

XX 25-MAR-2003 (revised)

DT 01-AUG-1995 (first entry)

XX Polypeptide encoded by hepatitis C virus Core/E1 sequence.

DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
XX  
KW classification; immunisation; prophylaxis; serotyping.

XX  
OS Hepatitis C virus type 3a.

XX  
PN WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

DR N-PSDB; AAQ78036.

XX New polynucleotide sequences from hepatitis C virus - and related  
PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
PT treatment, diagnosis and typing of HCV isolates.

PS - Claim 11; Page 118-119; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
CC nucleotides selected from an HCV type 3 genomic sequence, more  
CC particularly (i) the region spanning positions 417-957 of the Core/E1  
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated BR33-1-20.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ . Sequence 180 AA;

Query Match 98.7%; Score 951; DB 2; Length 180;  
 Best Local Similarity 97.8%; Pred. No. 1.2e-93;  
 Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAAGLEWRNTSG 60  
 QY 61 LVVLTNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCMTPTVAVKYVGATTASIR 120  
 DB 61 LVVLTNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCMTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
 DB 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

RESULT 10  
 AAR63283  
 ID AAR63283 standard; protein; 180 AA.  
 XX  
 AC AAR63283;  
 DT 25-MAR-2003 (revised)  
 DT 01-AUG-1995 (first entry)  
 XX  
 DE Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
 XX  
 KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
 KW classification; immunisation; prophylaxis; serotyping.  
 XX  
 OS Hepatitis C virus type 3a.  
 XX  
 FN WO9425601-A2.  
 XX  
 PD 10-NOV-1994.  
 XX  
 PF 27-APR-1994; 94WO-EP001323.  
 XX  
 PR 27-APR-1993; 93EP-00401099.  
 PR 05-AUG-1993; 93EP-00402019.  
 XX  
 PA (INNO-) INNOGENETICS NV SA.  
 XX  
 PI Maertens G, Stuyver L;  
 XX  
 DR WPI; 1994-358277/44.  
 DR N-PSDB; AAQ78035.

New polynucleotide sequences from hepatitis C virus - and related  
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
 PT treatment, diagnosis and typing of HCV isolates.  
 XX  
 PS Claim 11; Page 116-117; 404pp; English.  
 XX  
 CC Compositions comprising at least 5, and pref. 8 or more contiguous  
 CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-957 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated BR33-1-19.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ . Sequence 180 AA;

Query Match 98.7%; Score 951; DB 2; Length 180;  
 Best Local Similarity 97.8%; Pred. No. 1.2e-93;  
 Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAAGLEWRNTSG 60  
 QY 61 LVVLTNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCMTPTVAVKYVGATTASIR 120  
 DB 61 LVVLTNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCMTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
 DB 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

RESULT 11  
 AAR63279  
 ID AAR63279 standard; protein; 180 AA.  
 XX  
 AC AAR63279;  
 DT 25-MAR-2003 (revised)  
 DT 01-AUG-1995 (first entry)  
 XX

DE Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
 XX  
 KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
 KW classification; immunisation; prophylaxis; serotyping.  
 XX

OS Hepatitis C virus type 3a.  
 XX  
 FN WO9425601-A2.  
 XX  
 PD 10-NOV-1994.  
 XX  
 PF 27-APR-1994; 94WO-EP001323.  
 XX  
 PR 27-APR-1993; 93EP-00401099.  
 PR 05-AUG-1993; 93EP-00402019.  
 XX  
 PA (INNO-) INNOGENETICS NV SA.  
 XX  
 PI Maertens G, Stuyver L;  
 XX  
 DR WPI; 1994-358277/44.  
 DR N-PSDB; AAQ78031.

New polynucleotide sequences from hepatitis C virus - and related  
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
 PT treatment, diagnosis and typing of HCV isolates.  
 XX

PS Claim 11; Page 108; 404pp; English.  
 XX  
 CC Compositions comprising at least 5, and pref. 8 or more contiguous  
 CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-957 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
CC for immunisation against HCV, for the detection of antibodies directed  
CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
CC region of HCV subtype 3a and is taken from a clone designated HD10-2-21.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX  
SQ Sequence 180 AA;  
  
Query Match 98.3%; Score 948; DB 2; Length 180;  
Best Local Similarity 97.2%; Pred. No. 2.5e-93;  
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 VGAPVGGVARALAHGVRALGEGINPFCNLPFCGCSFIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALGEGINPFCNLPFCGCSFIFLLALFSLIHPAASLEWRNTSG 60  
  
QY 61 LVVLNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKVGATTASIR 120  
DB 61 LVVLNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKVGATTASIR 120  
  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFPRPRRHQTQVTCNCSLYPGHLSGHRMA 180  
DB 121 RHVDLLVGAATMCSALYVGMCGAVFLVGQAFPRPRRHQTQVTCNCSLYPGHLSGHRMA 180

RESULT 12  
AAR63277  
ID AAR63277 standard; protein; 180 AA.  
XX  
AC AAR63277;  
XX  
DT 25-MAR-2003 (revised)  
DT 20-JUL-1995 (first entry)  
XX  
XX Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
XX  
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
KW classification; immunisation; prophylaxis; serotyping.  
XX  
OS Hepatitis C virus type 3a.  
XX  
XX WO9425601-A2.  
XX  
PD 10-NOV-1994.  
XX  
XX 27-APR-1994; 94WO-EP001323.  
XX  
XX 27-APR-1993; 93EP-00401099.  
PR 05-AUG-1993; 93EP-00402019.  
XX  
XX (INNO-) INNOGENETICS NV SA.  
PA  
XX  
XX Maertens G, Stuyver L;  
PI  
XX  
XX WPI; 1994-358277/44.  
DR N-PSDB; AAQ78029.  
XX  
XX New polynucleotide sequences from hepatitis C virus - and related  
PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
PT treatment, diagnosis and typing of HCV isolates.  
XX  
XX - Claim 11; Page 104; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
CC nucleotides selected from an HCV type 3 genomic sequence, more  
CC particularly (i) the region spanning positions 417-957 of the Core/E1  
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
CC for immunisation against HCV, for the detection of antibodies directed  
CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
CC region of HCV subtype 3a and is taken from a clone designated HD10-2-5.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX  
SQ Sequence 180 AA;  
  
Query Match 98.3%; Score 948; DB 2; Length 180;  
Best Local Similarity 97.2%; Pred. No. 2.5e-93;  
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 VGAPVGGVARALAHGVRALGEGINPFCNLPFCGCSFIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALGEGINPFCNLPFCGCSFIFLLALFSLIHPAASLEWRNTSG 60  
  
QY 61 LVVLNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKVGATTASIR 120  
DB 61 LVVLNDCSNSIIVYEADDDVILHTPGCIPCVQDGNSTCWTPTVPTVAVKVGATTASIR 120  
  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFPRPRRHQTQVTCNCSLYPGHLSGHRMA 180  
DB 121 RHVDLLVGAATMCSALYVGMCGAVFLVGQAFPRPRRHQTQVTCNCSLYPGHLSGHRMA 180

RESULT 13  
AAR63278  
ID AAR63278 standard; protein; 180 AA.  
XX  
AC AAR63278;  
XX  
DT 25-MAR-2003 (revised)  
DT 21-JUL-1995 (first entry)  
XX  
XX Polypeptide encoded by hepatitis C virus Core/E1 sequence.  
XX  
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
KW classification; immunisation; prophylaxis; serotyping.  
XX  
OS Hepatitis C virus type 3a.  
XX  
XX WO9425601-A2.  
XX  
PD 10-NOV-1994.  
XX  
XX 27-APR-1994; 94WO-EP001323.  
XX  
XX 27-APR-1993; 93EP-00401099.  
PR 05-AUG-1993; 93EP-00402019.  
XX  
XX (INNO-) INNOGENETICS NV SA.  
PA  
XX  
XX Maertens G, Stuyver L;  
PI  
XX  
XX WPI; 1994-358277/44.  
DR N-PSDB; AAQ78030.  
XX  
XX New polynucleotide sequences from hepatitis C virus - and related  
PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
PT treatment, diagnosis and typing of HCV isolates.  
XX  
XX - Claim 11; Page 104; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
CC nucleotides selected from an HCV type 3 genomic sequence, more  
CC particularly (i) the region spanning positions 417-957 of the Core/E1  
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
CC as a probe for specific detection/classification of nucleic acid.

CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated HD10-2-14.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX

SQ Sequence 180 AA;

Query Match 97.6%; Score 941; DB 2; Length 180;  
 Best Local Similarity 96.7%; Pred. No. 1.4e-92;  
 Matches 174; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCIHPAASLEWNTSG 60  
 DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCIHPAASLEWNTSG 60  
 QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 DB 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVOTCNCSLYPGHLSGHRMA 180  
 DB 121 RHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVOTCNCSLYPGHLSGHRMA 180

RESULT 14  
 AAR94462  
 ID AAR94462 standard; protein; 3023 AA.

XX AAR94462;

DT 20-SEP-1996 (first entry)

DE Hepatitis C virus polypeptide.

XX hepatitis C virus; antibody; detection; diagnosis; vaccine; classify;  
 KW subtype.

OS Hepatitis C virus.

XX Key Location/Qualifiers  
 FH 1505..1520

FT Peptide /note= "this part of the sequence is missing from the  
 specification"

FT Peptide /note= "this part of the sequence is missing from the  
 specification"

XX JPO8056672-A.

XX 05-MAR-1996.

XX 26-AUG-1994; 94JP-00223933.

XX 26-AUG-1994; 94JP-00223933.

XX (SAYA/) SAYAMA K.

XX WPI; 1996-182301/19.  
 DR N-PSDB; AAT13279.

XX Hepatitis C virus genomic RNA, DNA and related proteins - useful for  
 PT detection, diagnosis and identification of hepatitis C virus sub-type.

XX Claim 4; Page 16-23; 25pp; Japanese.

XX The present sequence is a polypeptide comprising a 3023 amino acid  
 CC sequence which is encoded by a cDNA (AAT13279) to a hepatitis C virus  
 CC (HCV) genomic RNA. The protein can be easily detected by antibodies in an  
 CC assay for the detection of HCV. The DNA and the protein are useful for  
 CC classifying the subtype of HCV. At least a part of the protein may be  
 CC used as a vaccine against HCV

SQ Sequence 3023 AA;

Query Match 81.4%; Score 784.5; DB 2; Length 3023;  
 Best Local Similarity 79.0%; Pred. No. 3.5e-74;  
 Matches 143; Conservative 21; Mismatches 16; Indels 1; Gaps 1;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCIHPAASLEWNTSG 60  
 DB 140 IGAPVGGVARALAHGVRALEDGNYATGNLPGCSFSIFLLALFSLCIHPAASLEWNTSG 199  
 QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQ-DGNTSTCTWPTPTVAVKYVGATTASI 119  
 DB 200 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQVETDNNNTSCWTPISPTVAVKHPGVTTASI 259  
 QY 120 RSHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVOTCNCSLYPGHLSGHRM 179  
 DB 260 RHNVMNLVAPPTLCSALYVEDAFGAVSLVGQAFTFRPRQHKTQVTCNCSLYPGHSGHRM 319  
 QY 180 A 180  
 DB 320 A 320

RESULT 15

AAR63365  
 ID AAR63365 standard; protein; 193 AA.

XX AAR63365;

XX 25-MAR-2003 (revised)

DT 15-AUG-1995 (first entry)

DE Hepatitis C virus E1 polypeptide.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;  
 KW classification; immunisation; prophylaxis; serotyping.

OS Hepatitis C virus.

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

XX 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

XX N-PSDB; AAQ78103.

XX New polynucleotide sequences from hepatitis C virus - and related  
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
 PT treatment, diagnosis and typing of HCV isolates.

XX Disclosure; Page 225-226; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
 CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-857 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4832-  
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.  
 CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed

CC against HCV and for serotyping. This sequence corresponds to the E1  
CC region of HCV. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 193 AA;  
  
Query Match 78.1%; Score 753; DB 2; Length 193;  
Best Local Similarity 75.0%; Pred. No. 2.4e-72;  
Matches 135; Conservative 19; Mismatches 26; Indels 0; Gaps 0;  
  
QY 1 VGAPVGGVARALAGVRALEDGINFATGNLPGCSFSIFLLALFSLTHPAASLEWRNTSG 60  
Db 14 VGAPVGGVARALAGVRAVEDGINYATGNLPGCSFSIFLLALLSCLTVPASAQHYRNISG 73  
  
QY 61 LYVLTNDCSSSIYVEADVDVILHTPGCIPCVCQDGNSTCTWPTPTVAVKYVGATTASIR 120  
Db 74 IYHVTNDCPNSSIYVEADHHIMHLPQVPCVRTGNTSRCWVPLTPTVAAPYVGAPLESMR 133  
  
QY 121 SHVDLLVGAATMCSALYVGDMSGAVFLVGOAFTPRPRHQTVQTCNCSLYPGHLSHRMA 180  
Db 134 RHDLMVGAATVCSALYIGDLCGGVFLVGQMFTRPRRHWTQDCNCSIYDGHITGHRMA 193  
  
Search completed: December 1, 2004, 23:40:54  
Job time : 155 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:45:01 ; Search time 38 Seconds  
(without alignments)  
455.764 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 180  
Sequence: 1 VGAPVGGVARALAHGVRALE.....TVQTCNCSLYPGHLSGHRMA 180

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR\_79:.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	100.0	180	2	PC1304
2	87	48.3	180	2	PC1303
3	87	48.3	492	2	S41288
4	70	38.9	180	2	PC1305
5	66	36.7	411	2	PC2061
6	47	26.1	411	2	PC2060
7	19	10.6	874	2	JQ0881
8	19	10.6	3033	1	JQ1303
9	17	9.4	178	2	PS0388
10	17	9.4	189	2	S32740
11	17	9.4	315	2	PS0165
12	17	9.4	322	2	JN0265
13	17	9.4	369	2	S21471
14	17	9.4	441	2	S12707
15	17	9.4	513	2	PC1284
16	17	9.4	513	2	A44150
17	17	9.4	550	2	JH0711
18	17	9.4	640	2	JQ1584
19	17	9.4	782	2	S19876
20	17	9.4	782	2	S18032
21	17	9.4	782	2	S18031
22	17	9.4	782	2	S19875
23	17	9.4	787	2	PK0677
24	17	9.4	874	2	JQ0883
25	17	9.4	3010	1	A45573
26	17	9.4	3010	1	GNWVCJ
27	17	9.4	3010	1	GNWVTC
28	17	9.4	3010	1	GNWVTW
29	17	9.4	3010	1	S18030

#### ALIGNMENTS

##### RESULT 1

PC1304

genome polyprotein core/EI region (isolate BR36-9) - hepatitis C virus (strain type 3) (1)  
C;Species: hepatitis C virus  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: PC1304  
R;Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.  
Biochem. Biophys. Res. Commun. 192, 635-641, 1993  
A;Title: Analysis of the putative EI envelope and NS4a epitope regions of HCV type 3.  
A;Reference number: PC1300; MUID:93249436; PMID:7683463  
A;Accession: PC1304  
A;Molecule type: mRNA  
A;Residues: 1-180 <STU>  
A;Cross-references: UNIPROT:Q81278; DDBJ:D14599; NID:G303578; PIDN:BAA03448.1; PID:G3035;  
C;Experimental source: blood  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p  
F:53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status predicted  
F:125-154/Domain: transmembrane #status predicted <TM>  
F:57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 180; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 4.8e-177;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VGAPVGGVARALAHGVRALE	GINFATGNLPCGCSFIFLLALFSLIHPAASLEWNTSG	60
DB	1	VGAPVGGVARALAHGVRALE	GINFATGNLPCGCSFIFLLALFSLIHPAASLEWNTSG	60
QY	61	LYVLNDCSN	SIYVEADDDVILHTPGCIPCVQDGNSTCWTPTPTVAVKYVGATTASIR	120
DB	61	LYVLNDCSN	SIYVEADDDVILHTPGCIPCVQDGNSTCWTPTPTVAVKYVGATTASIR	120
QY	121	SHVDLLVGAATWCSALYVGD	MCGAVFLYVQAFTRPRRHQTVCNCSLYPGHLSGHRMA	180
DB	121	SHVDLLVGAATWCSALYVGD	MCGAVFLYVQAFTRPRRHQTVCNCSLYPGHLSGHRMA	180

##### RESULT 2

PC1303

genome polyprotein core/EI region (isolate HD10-2) - hepatitis C virus (strain type 3) (1)  
C;Species: hepatitis C virus  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: PC1303  
R;Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.  
Biochem. Biophys. Res. Commun. 192, 635-641, 1993  
A;Title: Analysis of the putative EI envelope and NS4a epitope regions of HCV type 3.  
A;Reference number: PC1300; MUID:93249436; PMID:7683463  
A;Accession: PC1303  
A;Molecule type: mRNA  
A;Residues: 1-180 <STU>  
A;Cross-references: UNIPROT:Q81279; DDBJ:D14603; NID:G303580; PIDN:BAA03452.1; PID:G30358

```

A;Experimental source: blood
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p
F;53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status predicted
F;125-154/Domain: transmembrane #status predicted <TM>
F;57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 48.3%; Score 87; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.7e-81;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFISFILALPSCLIHPAASLEWNTSG 60
DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFISFILALPSCLIHPAASLEWNTSG 60

QY 61 LYVLTNDCSNSSIVYEADVVILHTPGC 87
DB 61 LYVLTNDCSNSSIVYEADVVILHTPGC 87

RESULT 3
S41288
Genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S41288
R;Seelig, R.
submitted to the EMBL Data Library, December 1993
A;Reference number: S41288
A;Accession: S41288
A;Molecule type: genomic RNA
A;Residues: 1-492 <SEE>
A;Cross-references: UNIPROT:Q68870; EMBL:X76918
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; envelope protein; nonstructural protein; poly
F;1-191/Product: core protein #status predicted <COR>
F;192-372/Product: envelope protein #status predicted <ENV>
F;373-492/Product: NS1 protein (fragment) #status predicted <NS1>

Query Match 48.3%; Score 87; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 3.9e-81;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFISFILALPSCLIHPAASLEWNTSG 60
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFISFILALPSCLIHPAASLEWNTSG 199

QY 61 LYVLTNDCSNSSIVYEADVVILHTPGC 87
DB 200 LYVLTNDCSNSSIVYEADVVILHTPGC 226

RESULT 4
PC1305
Genome polyprotein core/E1 region (isolate BR33-1) - hepatitis C virus (strain type 3)
C;Species: hepatitis C virus
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: PC1305
R;Stuyver, L.; Van Arnhem, W.; Wyseur, A.; DeLeys, R.; Maertens, G.
Biochem. Biophys. Res. Commun. 192, 635-641, 1993
A;Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.
A;Reference number: PC1300; MUID:93249436; PMID:7683463
A;Accession: PC1305
A;Molecule type: mRNA
A;Residues: 1-180 <STU>
A;Cross-references: UNIPROT:Q81277; DBJ:DJ14596; NID:G303576; PIDN:BAA03445.1; PID:G3035
A;Experimental source: blood
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p
F;53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status predicted
F;125-154/Domain: transmembrane #status predicted <TM>
F;57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match 38.9%; Score 70; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 4.9e-64;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 YVGATTASIRSHVDLLVGAATWCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLY 170
DB 111 YVGATTASIRSHVDLLVGAATWCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLY 170
QY 171 PGHLSGHERMA 180
DB 171 PGHLSGHERMA 180

RESULT 5
PC2061
Genome polyprotein N2 - hepatitis C virus
N;Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: PC2061
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A;Title: Identification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MUID:94197744; PMID:8147893
A;Accession: PC2061
A;Molecule type: mRNA
A;Residues: 1-411 <LIJ>
A;Cross-references: UNIPROT:Q81813; GB:L12355; NID:9410169; PIDN:AAA20155.1; PID:9410170
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein; i
F;192-383/Product: envelope protein E1 #status predicted <SP>
F;384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>
F;196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.7%; Score 66; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.3e-59;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 TTASIRSHVDLLVGAATWCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHL 174
DB 254 TTASIRSHVDLLVGAATWCSALYVGMCGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHL 313
QY 175 SGRHMA 180
DB 314 SGRHMA 319

RESULT 6
PC2060
Genome polyprotein N1 - hepatitis C virus
N;Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: PC2060
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A;Title: Identification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MUID:94197744; PMID:8147893
A;Accession: PC2060
A;Molecule type: mRNA
A;Residues: 1-411 <LIJ>
A;Cross-references: UNIPROT:Q81489
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein; i
F;192-383/Product: envelope protein E1 #status predicted <SP>
F;384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>
F;196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.18%; Score 47; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.1e-40;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLI 47  
140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLI 186

Db

RESULT 7

JQ0881

genome polyprotein - hepatitis C virus (strain J6) (fragments)

N;Contains: NS5 protein

C;Species: hepatitis C virus

C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jul-2004

C;Accession: JQ0881

R;Okamoto, H.

submitted to JIPID, January 1991

A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis C virus

A;Reference number: JQ0879

A;Accession: JQ0881

A;Molecule type: genomic RNA

A;Residues: 1-874 <OKA>

A;Cross-references: UNIPROT:Q71ZYS

A;Experimental source: strain J6

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: polyprotein

F;510-874/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match 10.6%; Score 19; DB 2; Length 874;  
Best Local Similarity 100.0%; Pred. No. 4.5e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NFATGNLPGCSFSIFLLAL 42  
163 NFATGNLPGCSFSIFLLAL 181

Db

RESULT 8

JQ1303

genome polyprotein - hepatitis C virus (isolate HC-J6)

N;Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C;Species: hepatitis C virus

C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C;Accession: JQ1303

R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.

J. Gen. Virol. 72, 2697-2704, 1991

A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human

A;Reference number: JQ1303; MUID:92044440; PMID:1656196

A;Accession: JQ1303

A;Molecule type: genomic RNA

A;Residues: 1-3033 <OKA>

A;Cross-references: UNIPROT:P26660; GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651

A;Experimental source: isolate HC-J6 from a Japanese individual

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane

F;116-191/Product: capsid protein C #status predicted <CPC>

F;192-389/Product: envelope protein M #status predicted <EPM>

F;390-733/Product: major envelope protein E #status predicted <MEE>

F;734-1010/Product: nonstructural protein NS1 #status predicted <NS1>

F;1011-1619/Product: nonstructural protein NS2 #status predicted <NS2>

F;1316-1321/Region: nucleotide-binding motif B

F;1320-1323/Region: DEXH motif

F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>

F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>

F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>

F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,28

Query Match 10.6%; Score 19; DB 1; Length 3033;  
Best Local Similarity 100.0%; Pred. No. 1.3e-10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NFATGNLPGCSFSIFLLAL 42  
163 NFATGNLPGCSFSIFLLAL 181

Db

RESULT 9

PS0388

genome polyprotein - hepatitis C virus (isolate GM1) (fragment)

N;Contains: amino end of envelope protein M; carboxyl end of capsid protein C

C;Species: hepatitis C virus

A;Note: host Homo sapiens (man)

C;Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 09-Jul-2004

C;Accession: PS0388

R;Fuchs, K.; Morz, M.; Schreier, E.; Zchoval, R.; Deinhardt, F.; Roggendorf, M.

Gene 103, 163-169, 1991

A;Title: Characterization of nucleotide sequences from European hepatitis C virus isolate

A;Reference number: UN0265; MUID:91365241; PMID:1653756

A;Accession: PS0388

A;Molecule type: genomic RNA

A;Residues: 1-178 <FUC>

A;Cross-references: UNIPROT:Q81275; GB:M61719; NID:G329757; PIDN:AAA45536.1; PID:G387654

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; envelope protein; polyprotein; transmembrane protein

F;1-109/Product: capsid protein C (fragment) #status predicted <CPC>

F;110-178/Product: envelope protein M (fragment) #status predicted <EPM>

F;163-178/Domain: transmembrane #status predicted <TM1>

Query Match 9.4%; Score 17; DB 2; Length 178;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42  
159 ATGNLPGCSFSIFLLAL 175

Db

RESULT 10

S32740

polyprotein - hepatitis C virus (isolate Russian) (fragment)

N;Contains: capsid protein C; envelope protein M

C;Species: hepatitis C virus

C;Date: 15-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004

C;Accession: S32740

R;Vassilev, V.B.; Viazov, S.O.; Kotova, E.Y.; Nosikov, V.V.

submitted to the EMBL Data Library, April 1993

A;Description: Evidence of new HCV variant of European isolate in Russia.

A;Reference number: S32740

A;Accession: S32740

A;Molecule type: genomic RNA

A;Residues: 1-189 <VAS>

A;Cross-references: UNIPROT:Q68873; EMBL:X71407

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; envelope protein; polyprotein

F;1-115/Product: capsid protein C #status predicted <CPC>

F;116-189/Product: envelope protein M #status predicted <EPM>

Query Match 9.4%; Score 17; DB 2; Length 189;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42  
165 ATGNLPGCSFSIFLLAL 181

Db

RESULT 11

PS0165

envelope glycoprotein (clone 168) - hepatitis C virus (fragment)

C;Species: hepatitis C virus

C;Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 09-Jul-2004

C;Accession: PS0165

R;Hajikata, M.; Kato, N.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Shimotohno, K.

Biochem. Biophys. Res. Commun. 175, 220-228, 1991

A;Title: Hypervariable regions in the putative glycoprotein of hepatitis C virus.

A;Reference number: PN0011; MUID:91151533; PMID:1847805

A;Accession: PS0165

```

A:Molecule type: genomic RNA
A:Residues: 1-315 <H1U>
A:Cross-references: UNIPROT:Q00906; GB:D00691
A>Note: the authors translated the codon GAC for residues 27 and 112 as Glu and CTT for
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: glycoprotein; polypeptide
F:200-209/Region: hypervariable 1 #status predicted
F:283-289/Region: hypervariable 2 #status predicted
F:15,18,59,114,134,226,232,239,257/Binding site: carbohydrate (Asn) (covalent) #status ph

Query Match          9.4%; Score 17; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TNDGNSNSIVYEADDDVI 81
DB 13 TNDGNSNSIVYEADDDVI 29

RESULT 12
JN0265
genome polypeptide - hepatitis C virus (isolate GM2) (fragments)
N:Contains: amino end of envelope protein M; carboxyl end of capsid protein C; fragment
C:Species: hepatitis C virus
A>Note: host Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C:Accession: JN0265
R:Fuchs, K.; Motz, M.; Schreier, E.; Zachoal, R.; Deinhardt, F.; Roggendorf, M.
Gene 103, 163-169, 1991
A>Title: Characterization of nucleotide sequences from European hepatitis C virus isolat
A:Reference number: JN0265; MUID:91365241; PMID:1653756
A:Accession: JN0265
A:Molecule type: genomic RNA
A:Residues: 1-322 <FUG>
A:Cross-references: UNIPROT:Q03729; GB:M61717; GB:M61718
A>Note: the authors translated the codon ACA for residue 198 as Tyr
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; envelope protein; glycoprotein; polypeptide; transmembrane p
F:1-109/Product: capsid protein C (fragment) #status predicted <COR>
F:110-178/Product: envelope protein M (fragment) #status predicted <BPM>
F:163-178/Domain: transmembrane #status predicted <TML>
F:179-322/Product: major envelope protein E (fragment) #status predicted <ENV>
F:263-269/Domain: transmembrane #status predicted <TM>
F:191,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          9.4%; Score 17; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42
DB 159 ATGNLPGCSFSIFLLAL 175

RESULT 13
S21471
genome polypeptide - hepatitis C virus (fragment)
N:Contains: capsid protein; envelope protein
C:Species: hepatitis C virus
C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S21471
R:McGon, W.K.
submitted to the EMBL Data Library, April 1992
A:Reference number: S21471
A:Accession: S21471
A:Molecule type: genomic RNA
A:Residues: 1-369 <MOG>
A:Cross-references: UNIPROT:Q68869; EMBL:X65924; NID:g59466; PIDN:CAA46717.1; PID:g59467
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; envelope protein; glycoprotein; polypeptide

Query Match          9.4%; Score 17; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;

```

```

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42
DB 165 ATGNLPGCSFSIFLLAL 181

RESULT 14
S12707
genome polypeptide - hepatitis C virus (fragment)
N:Contains: core protein; envelope protein
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S12707
R:Takeuchi, K.; Kubo, Y.; Boonmar, S.; Watanabe, Y.; Katayama, T.; Choo, Q.L.; Kuo, G.; F
Nucleic Acids Res. 18, 4626, 1990
A>Title: Nucleotide sequence of core and envelope genes of the hepatitis C virus genome
A:Reference number: S12707; MUID:90356432; PMID:2117749
A:Accession: S12707
A:Molecule type: genomic RNA
A:Residues: 1-441 <TAK>
A:Cross-references: UNIPROT:Q81776; EMBL:D00574; NID:g221656; PIDN:BAA00452.1; PID:g22166
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: polypeptide

Query Match          9.4%; Score 17; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42
DB 165 ATGNLPGCSFSIFLLAL 181

RESULT 15
PC1284
genome polypeptide - hepatitis C virus (isolate HC-J4) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1284
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A>Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1284
A:Molecule type: genomic RNA
A:Residues: 1-513 <OKA>
A:Cross-references: UNIPROT:Q81221; GB:D00832; NID:g221513; PIDN:BAA00706.1; PID:g221514
C:Superfamily: hepatitis C virus genome polypeptide

Query Match          9.4%; Score 17; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ATGNLPGCSFSIFLLAL 42
DB 165 ATGNLPGCSFSIFLLAL 181

Search completed: December 1, 2004, 23:54:52
Job time : 39 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:48:17 ; Search time 142 Seconds  
(without alignments)  
452.057 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 180

Sequence: 1 VGAPVGVGVARALAHGVRALE.....TVQTNCNSLYPHLSGRHMA 180

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1582122 seqs, 356623098 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1582122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	100.0	180	10	US-09-899-046-20
2	180	100.0	180	10	US-09-899-046-20
3	180	100.0	180	10	US-09-899-046-22
4	180	100.0	180	10	US-09-878-281-20
5	180	100.0	180	10	US-09-878-281-22
6	180	100.0	180	10	US-09-878-281-24
7	180	100.0	180	10	US-09-873-224-20
8	180	100.0	180	10	US-09-873-224-22
9	180	100.0	208	9	US-09-973-025-30
10	180	100.0	208	10	US-09-899-303-30
11	180	100.0	208	10	US-09-995-808-30
12	180	100.0	208	10	US-09-995-860-30
13	180	100.0	208	10	US-09-935-791-30
14	180	100.0	208	16	US-10-321-798-30
15	180	100.0	209	14	US-10-128-590-91

14	180	100.0	209	14	US-10-128-587A-91	Sequence 91, Appl
15	22	51.1	187	15	US-10-651-165-209	Sequence 209, App
16	87	48.3	180	10	US-09-899-046-14	Sequence 14, Appl
17	87	48.3	180	10	US-09-899-046-18	Sequence 18, Appl
18	87	48.3	180	10	US-09-878-281-14	Sequence 14, Appl
19	87	48.3	180	10	US-09-878-281-18	Sequence 18, Appl
20	87	48.3	180	10	US-09-873-224-14	Sequence 14, Appl
21	87	48.3	180	10	US-09-873-224-18	Sequence 18, Appl
22	87	48.3	187	15	US-10-651-165-207	Sequence 207, App
23	87	48.3	193	15	US-10-651-165-210	Sequence 210, App
24	70	38.9	180	10	US-09-899-046-24	Sequence 24, Appl
25	70	38.9	180	10	US-09-899-046-26	Sequence 26, Appl
26	70	38.9	180	10	US-09-899-046-28	Sequence 28, Appl
27	70	38.9	180	10	US-09-878-281-24	Sequence 24, Appl
28	70	38.9	180	10	US-09-878-281-26	Sequence 26, Appl
29	70	38.9	180	10	US-09-878-281-28	Sequence 28, Appl
30	70	38.9	180	10	US-09-873-224-24	Sequence 24, Appl
31	70	38.9	180	10	US-09-873-224-26	Sequence 26, Appl
32	70	38.9	180	10	US-09-873-224-28	Sequence 28, Appl
33	70	38.9	187	15	US-10-651-165-208	Sequence 208, App
34	55	30.6	180	10	US-09-899-046-16	Sequence 16, Appl
35	55	30.6	180	10	US-09-878-281-16	Sequence 16, Appl
36	55	30.6	180	10	US-09-873-224-16	Sequence 16, Appl
37	51	28.3	189	15	US-10-450-649-9	Sequence 9, Appli
38	35	19.4	35	14	US-10-318-200-35	Sequence 35, Appl
39	35	19.4	35	16	US-10-685-435-5	Sequence 5, Appl
40	26	14.4	149	9	US-09-851-138-52	Sequence 52, Appl
41	26	14.4	149	15	US-10-651-165-284	Sequence 284, App
42	24	13.3	24	16	US-10-685-435-6	Sequence 6, Appli
43	22	12.2	149	9	US-09-851-138-30	Sequence 30, Appl
44	22	12.2	149	15	US-10-651-165-283	Sequence 283, App
45	21	11.7	319	15	US-10-651-165-211	Sequence 211, App

ALIGNMENTS

RESULT 1

US-09-899-046-20  
; Sequence 20, Application US/09899046  
; Publication No. US2003008274A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/899,046  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-899-046-20

Query Match 100.0%; Score 180; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.5e-164;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGAPVGVGVARALAHGVRALEDGINFATGNLPCGSPSIFLLALFSCLIHPAALEWRNTSG 60  
Db 1 VGAPVGVGVARALAHGVRALEDGINFATGNLPCGSPSIFLLALFSCLIHPAALEWRNTSG 60

QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
Db 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 2

US-09-899-046-22  
; Sequence 22, Application US/09899046  
; Publication No. US2003008274A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/899,046  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-899-046-22

Query Match 100.0%; Score 180; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.5e-164;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
Db 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 3

US-09-878-281-20  
; Sequence 20, Application US/09878281  
; Publication No. US20030032005A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/878,281  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455

; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-878-281-20

Query Match 100.0%; Score 180; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.5e-164;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
Db 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 4

US-09-878-281-22  
; Sequence 22, Application US/09878281  
; Publication No. US20030032005A1  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: New sequences of hepatitis C virus  
; NUMBER OF SEQUENCES: 270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/878,281  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/362,455  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-878-281-22

Query Match 100.0%; Score 180; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.5e-164;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSCLIHPAASLEWNTSG 60  
QY 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
Db 61 LYVLTNDCNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120  
QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHRMA 180

## RESULT 5

US-09-873-224-20

```
; Sequence 20, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics sa.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-873-224-20
Query Match 100.0%; Score 180; DB 10; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.5e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
Qy 61 LVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Db 61 LVLTNDCSNSIYVEADDVILHTPGICPCVQDGNSTCWTPTVAVKYVGATTASIR 120
Qy 121 SHVDLLVGAATMCSALYVGDVMDGAVFLVGQAFTRPRRHQTVCNCNCSLYPGHLSGHRMA 180
Db 121 SHVDLLVGAATMCSALYVGDVMDGAVFLVGQAFTRPRRHQTVCNCNCSLYPGHLSGHRMA 180
RESULT 7
US-09-873-025-30
; Sequence 30, Application US/09973025
; Publication No. US20020182706A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; BOSMAN, FONS
; DE MARTYNOFF, GUY
; BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/973,025
; FILING DATE: 10-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
```

```
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-973-025-30

Query Match 100.0%; Score 180; DB 9; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 60
Db 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 81

QY 61 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120
Db 82 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 8
US-09-899-303-30
; Sequence 30, Application US/09899303
; Publication No. US20030036110A1
; GENERAL INFORMATION:
; APPLICANT: BOSMAN, FONS
; DE MARTYNOFF, GUY
; BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,303
; FILING DATE: 06-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-899-303-30

Query Match 100.0%; Score 180; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 60
Db 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 81

QY 61 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120
Db 82 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 9
US-09-995-808-30
; Sequence 30, Application US/09995808
; Publication No. US20030095980A1
; GENERAL INFORMATION:
; APPLICANT: Immunogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-70
; CURRENT APPLICATION NUMBER: US/09/995,808
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 30
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-995-808-30

Query Match 100.0%; Score 180; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 60
Db 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCLIHHPAASLEWNTSG 81

QY 61 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 120
Db 82 LYVLTDSCNSSIYVEADDDVILHTPGCIPCVQDGNSTCMTPTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 180
Db 142 SHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTRPRRHQTVQTCNCSLYPGHLSGHRMA 201

RESULT 10
US-09-995-860-30
; Sequence 30, Application US/09995860
; Publication No. US20030118603A1
; GENERAL INFORMATION:
; APPLICANT: Immunogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-69
; CURRENT APPLICATION NUMBER: US/09/995,860
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 30
; LENGTH: 208
; TYPE: PRT
```

```
; ORGANISM: Hepatitis C virus
US-09-995-860-30

Query Match      100.0%; Score 180; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
DB 82 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 180
DB 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 201

RESULT 11
US-09-995-791-30
; Sequence 30, Application US/09995791
; Publication No. US20030147918A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-68
; CURRENT APPLICATION NUMBER: US/09/995,791
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent in 3.1
; SEQ ID NO 30
; TYPE: PRT
; LENGTH: 208
; ORGANISM: Hepatitis C virus
US-09-995-791-30

Query Match      100.0%; Score 180; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
DB 82 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 180
DB 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 201

RESULT 12
US-10-321-798-30
; Sequence 30, Application US/10321798
; Publication No. US20040126395A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-93
; CURRENT APPLICATION NUMBER: US/10/321,798
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 60/418,359
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 10/020,510
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent in 3.1
```

```
; SEQ ID NO 30
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-321-798-30

Query Match      100.0%; Score 180; DB 16; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 81

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
DB 82 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 180
DB 142 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 201

RESULT 13
US-10-128-590-91
; Sequence 91, Application US/10128590
; Publication No. US20030108561A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Core-glycosylated HCV envelope proteins
; FILE REFERENCE: 135 PCT
; CURRENT APPLICATION NUMBER: US/10/128,590
; CURRENT FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 91
; LENGTH: 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-128-590-91

Query Match      100.0%; Score 180; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.7e-164;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 23 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 82

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 120
DB 83 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCWTPTVAVKYVGATTASIR 142

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 180
DB 143 SHVDLLVGAATMCSALYVGDGCGAVFLVGQAFTFRPRRHQVQTCNCSLYPGHLSGHRMA 202

RESULT 14
US-10-128-587A-91
; Sequence 91, Application US/10128587A
; Publication No. US20030152940A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Constructs and methods for expression of recombinant HCV envelope
; FILE REFERENCE: 134 PCT
; CURRENT APPLICATION NUMBER: US/10/128,587A
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 91
; LENGTH: 209
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:45:46 ; Search time 39 seconds  
(without alignments)  
306.083 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 180

Sequence: 1 VGAPVGGVARALAHGVRALE.....TVQTCNCSLYPGHLSGHRMA 180

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	100.0	180	4	US-09-878-281A-20
2	180	100.0	180	4	US-09-878-281A-22
3	180	100.0	208	3	US-08-612-973-30
4	180	100.0	208	3	US-08-927-597-30
5	92	51.1	187	4	US-08-635-886C-209
6	92	51.1	187	4	US-08-974-690C-209
7	87	48.3	180	4	US-09-878-281A-14
8	87	48.3	180	4	US-09-878-281A-18
9	87	48.3	187	4	US-08-635-886C-207
10	87	48.3	187	4	US-08-974-690C-207
11	87	48.3	193	4	US-08-635-886C-210
12	87	48.3	193	4	US-08-974-690C-210
13	70	38.9	180	4	US-09-878-281A-24
14	70	38.9	180	4	US-09-878-281A-26
15	70	38.9	180	4	US-09-878-281A-28
16	70	38.9	187	4	US-08-635-886C-208
17	70	38.9	187	4	US-08-974-690C-208
18	70	38.9	192	1	US-08-086-428B-87
19	70	38.9	192	1	US-08-086-428B-88
20	70	38.9	192	2	US-08-468-570-87
21	70	38.9	192	2	US-08-468-570-88
22	70	38.9	192	2	US-08-290-665A-87
23	70	38.9	192	2	US-08-290-665A-88
24	70	38.9	192	4	US-08-466-601A-87
25	70	38.9	192	4	US-08-466-601A-88
26	70	38.9	192	5	PCT-US95-10398-87
27	70	38.9	192	5	PCT-US95-10398-88

28 55 30.6 180 4 US-09-878-281A-16 Sequence 16, Appl  
29 52 28.9 191 2 US-08-290-665A-187 Sequence 187, App  
30 52 28.9 191 2 US-08-290-665A-189 Sequence 189, App  
31 52 28.9 191 2 US-08-290-665A-190 Sequence 190, App  
32 52 28.9 191 5 PCT-US95-10398-187 Sequence 187, App  
33 52 28.9 191 5 PCT-US95-10398-189 Sequence 189, App  
34 52 28.9 191 5 PCT-US95-10398-190 Sequence 190, App  
35 48 26.7 192 1 US-08-086-428B-90 Sequence 90, Appl  
36 48 26.7 192 2 US-08-468-570-90 Sequence 90, Appl  
37 48 26.7 192 2 US-08-290-665A-90 Sequence 90, Appl  
38 48 26.7 192 4 US-08-466-601A-90 Sequence 90, Appl  
39 48 26.7 192 5 PCT-US95-10398-90 Sequence 90, Appl  
40 46 25.6 191 2 US-08-290-665A-188 Sequence 188, App  
41 46 25.6 191 5 PCT-US95-10398-188 Sequence 188, App  
42 41 22.8 192 1 US-08-086-428B-89 Sequence 89, Appl  
43 41 22.8 192 2 US-08-468-570-89 Sequence 89, Appl  
44 41 22.8 192 2 US-08-290-665A-89 Sequence 89, Appl  
45 41 22.8 192 4 US-08-466-601A-89 Sequence 89, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-878-281A-20

; Sequence 20, Application US/09878281A

; Patent No. 6762024

; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.

; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy

; FILE REFERENCE: 35

; CURRENT APPLICATION NUMBER: US/09/878,281A

; CURRENT FILING DATE: 2001-06-12

; NUMBER OF SEQ ID NOS: 284

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20

; LENGTH: 180

; TYPE: PRT

; ORGANISM: hepatitis C virus

US-09-878-281A-20

Query Match 100.0%; Score 180; DB 4; Length 180;

Best Local Similarity 100.0%; Pred. No. 2e-165;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPCCSFSIFLLALFSLIHPAASLEWRNTSG 60

DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPCCSFSIFLLALFSLIHPAASLEWRNTSG 60

QY 61 LVYLTNDCSNSIVYEADDDVILHTPGCTPCVODGNTSTCWTPTVTVAKYVGATTASIR 120

DB 61 LVYLTNDCSNSIVYEADDDVILHTPGCTPCVODGNTSTCWTPTVTVAKYVGATTASIR 120

QY 121 SHVDLLVGAATMCSALYVGDMDCGAVFLVGOAFTFRFRHQTQTCNCSLYPGHLSGHRMA 180

DB 121 SHVDLLVGAATMCSALYVGDMDCGAVFLVGOAFTFRFRHQTQTCNCSLYPGHLSGHRMA 180

##### RESULT 2

US-09-878-281A-22

; Sequence 22, Application US/09878281A

; Patent No. 6762024

; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.

; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy

; FILE REFERENCE: 35

; CURRENT APPLICATION NUMBER: US/09/878,281A

; CURRENT FILING DATE: 2001-06-12

; NUMBER OF SEQ ID NOS: 284

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 22

```
; LENGTH: 180
; TYPE: PRT
; -ORGANISM: hepatitis C virus
US-09-878-281A-22

Query Match
Best Local Similarity 100.0%; Score 180; DB 4; Length 180;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 60
Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 60

QY 61 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 120
Db 61 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 120

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 180
Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 180

RESULT 3
US-08-612-973-30
; Sequence 30, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUJSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-612-973-30

Query Match
Best Local Similarity 100.0%; Score 180; DB 3; Length 208;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 60
Db 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 81

QY 61 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 120
Db 82 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 180
Db 142 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 201

RESULT 4
US-08-927-597-30
; Sequence 30, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUJSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 11-MAR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-927-597-30

Query Match
Best Local Similarity 100.0%; Score 180; DB 3; Length 208;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 60
Db 22 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFSIFLLALFSLIHPAASLEWNTSG 81

QY 61 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 120
Db 82 LYVLINDCSNSSIYVEADDVILHTPGCIPCVQDGNSTCMTPTVPTVAVKYVGATTASIR 141

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 180
Db 142 SHVDLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGRMA 201

RESULT 5
```

```
US-08-635-886C-209
; Sequence 209, Application US/08635886C
; Patent No. 655114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635.886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 209
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-209

Query Match      51.1%; Score 92; DB 4; Length 187;
Best Local Similarity 100.0%; Pred. No. 1e-80;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 PCVQDGNSTCTWPTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLV 148
Db 96 PCVQDGNSTCTWPTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLV 155

Qy 149 GQAFTFRPRRHQTQTCNCSLYPGHLSGRMA 180
Db 156 GQAFTFRPRRHQTQTCNCSLYPGHLSGRMA 187

RESULT 5
US-08-974-690C-209
; Sequence 209, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974.690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 209
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-209

Query Match      51.1%; Score 92; DB 4; Length 187;
Best Local Similarity 100.0%; Pred. No. 1e-80;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 PCVQDGNSTCTWPTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLV 148
Db 96 PCVQDGNSTCTWPTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLV 155

Qy 149 GQAFTFRPRRHQTQTCNCSLYPGHLSGRMA 180
Db 156 GQAFTFRPRRHQTQTCNCSLYPGHLSGRMA 187
```

## RESULT 7

```
US-09-878-281A-14
; Sequence 14, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878.281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-14
```

```
Query Match      48.3%; Score 87; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.4e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60
Db 1 VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60

Qy 61 LVYLTNDCSNSSIVYEADDDVILHTPGC 87
Db 61 LVYLTNDCSNSSIVYEADDDVILHTPGC 87
```

## RESULT 8

```
US-09-878-281A-18
; Sequence 18, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878.281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-18
```

```
Query Match      48.3%; Score 87; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.4e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60
Db 1 VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60

Qy 61 LVYLTNDCSNSSIVYEADDDVILHTPGC 87
Db 61 LVYLTNDCSNSSIVYEADDDVILHTPGC 87
```

## RESULT 9

```
US-08-635-886C-207
; Sequence 207, Application US/08635886C
; Patent No. 655114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
```

```
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 207
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-207

Query Match      48.3%; Score 87; DB 4; Length 187;
Best Local Similarity 100.0%; Pred. No. 6.6e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 60
Db 8 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 67

QY 61 LYVLTDNCSSSIYVEADDVILHTPGC 87
Db 68 LYVLTDNCSSSIYVEADDVILHTPGC 94

RESULT 10
US-08-974-690C-207
; Sequence 207, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 207
; LENGTH: 187
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-207

Query Match      48.3%; Score 87; DB 4; Length 187;
Best Local Similarity 100.0%; Pred. No. 6.6e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 60
Db 8 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 67

QY 61 LYVLTDNCSSSIYVEADDVILHTPGC 87
Db 68 LYVLTDNCSSSIYVEADDVILHTPGC 94

RESULT 11
US-08-635-886C-210
; Sequence 210, Application US/08635886C
; Patent No. 6555114
```

```
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; LENGTH: 193
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-210

Query Match      48.3%; Score 87; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 6.8e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 60
Db 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 73

QY 61 LYVLTDNCSSSIYVEADDVILHTPGC 87
Db 74 LYVLTDNCSSSIYVEADDVILHTPGC 100

RESULT 12
US-08-974-690C-210
; Sequence 210, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; LENGTH: 193
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-210

Query Match      48.3%; Score 87; DB 4; Length 193;
Best Local Similarity 100.0%; Pred. No. 6.8e-76;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 60
Db 14 VGAPVGGVARALAHGVRALEDGINFATGNLPGCCSFIFLLALFSLIHPAASLEWRNTSG 73

QY 61 LYVLTDNCSSSIYVEADDVILHTPGC 87
Db 74 LYVLTDNCSSSIYVEADDVILHTPGC 100

RESULT 13
US-08-974-690C-210
```

```
US-09-878-281A-24
; Sequence 24, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-24

Query Match      38.9%; Score 70; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.5e-59;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Db      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Qy      171 PGLSGHRMA 180
Db      171 PGLSGHRMA 180

RESULT 14
US-09-878-281A-26
; Sequence 26, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-26

Query Match      38.9%; Score 70; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.5e-59;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Db      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Qy      171 PGLSGHRMA 180
Db      171 PGLSGHRMA 180

RESULT 15
US-09-878-281A-28
; Sequence 28, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
```

```
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 180
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-28

Query Match      38.9%; Score 70; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.5e-59;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Db      111 YVGATTASIRSHVDLLVGAATMCSALYVGDMDGAVFLVGQAFTPRRRHQTVCNCSLY 170
Qy      171 PGLSGHRMA 180
Db      171 PGLSGHRMA 180

Search completed: December 1, 2004, 23:55:37
Job time : 40 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:27:29 ; Search time 193 Seconds  
(without alignments)  
536.619 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 964

Sequence: 1 VGAPVGGVARALAHGVRALE.....TVQTCNCSLYPGHLSGRMA 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	964	100.0	180	2	Q81278 hepatitis c
2	960	99.6	503	2	Q81752 hepatitis c
3	960	99.6	3021	2	Q68870 hepatitis c
4	960	99.6	3021	2	Q81258 hepatitis c
5	955	99.1	411	2	Q81813 hepatitis c
6	951	98.7	180	2	Q81277 hepatitis c
7	948	98.3	180	2	Q81279 hepatitis c
8	944	97.9	503	2	Q81489 hepatitis c
9	944	97.9	503	2	Q81827 hepatitis c
10	942	97.7	3021	2	Q82933 hepatitis c
11	935	97.0	503	2	Q81822 hepatitis c
12	933	96.8	3021	2	Q81495 hepatitis c
13	865	89.7	415	2	Q81546 hepatitis c
14	859	89.1	415	2	Q81554 hepatitis c
15	857	88.9	415	2	Q81552 hepatitis c
16	845	87.7	415	2	Q81266 hepatitis c
17	844	87.6	415	2	Q81550 hepatitis c
18	841	87.2	415	2	Q81548 hepatitis c
19	828	85.9	416	2	Q81265 hepatitis c
20	798	82.8	415	2	Q68806 hepatitis c
21	795	82.5	149	2	Q68728 hepatitis c
22	792	82.2	415	2	Q68794 hepatitis c
23	791	82.1	419	2	Q81345 hepatitis c
24	787	81.6	149	2	Q68738 hepatitis c
25	784.5	81.4	514	2	Q6XJ43 hepatitis c
26	784.5	81.4	514	2	Q6XJ45 hepatitis c
27	784.5	81.4	514	2	Q6XJ46 hepatitis c
28	784.5	81.4	514	2	AAP69954 hepatitis c
29	784.5	81.4	514	2	AAP69955 hepatitis c
30	784.5	81.4	514	2	AAP69957 hepatitis c
31	784.5	81.4	515	2	Q6XJ40 hepatitis c

32	784.5	81.4	515	2	Q6XJ41 hepatitis c
33	784.5	81.4	515	2	Q6XJ42 hepatitis c
34	784.5	81.4	515	2	Q6XJ44 hepatitis c
35	784.5	81.4	515	2	AAP69956 hepatitis c
36	784.5	81.4	515	2	AAP69958 hepatitis c
37	784.5	81.4	515	2	AAP69959 hepatitis c
38	784.5	81.4	515	2	AAP69960 hepatitis c
39	784.5	81.4	3023	2	Q81487 hepatitis c
40	783	81.2	415	2	Q68802 hepatitis c
41	778	80.7	415	2	Q68799 hepatitis c
42	778	80.7	3019	2	Q68801 hepatitis c
43	772	80.1	415	2	Q68808 hepatitis c
44	769	79.8	149	2	Q81282 hepatitis c
45	753	78.1	193	2	Q81218 hepatitis c

ALIGNMENTS

RESULT 1

Q81278	PRELIMINARY;	PRT;	180 AA.
ID	Q81278;		
AC	Q81278;		
DT	01-NOV-1996 (TREMELrel. 01, Created)		
DT	01-NOV-1996 (TREMELrel. 01, Last sequence update)		
DT	01-JUN-2003 (TREMELrel. 24, Last annotation update)		
DE	Core protein-envelope 1 protein (Fragment).		
GN	Name=core-E1;		
OS	Hepatitis C virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepacivirus.		
OX	NCBI_TaxID=11103;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93249436; PubMed=7683463;		
RA	Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;		
RT	"Analysis of the putative E1 envelope and NS4a epitope region of HCV type 3.";		
RL	Biochem. Biophys. Res. Commun. 192:635-641 (1993).		
DR	EMBL; D14599; BAA03448.1; -.		
DR	PIR; PC1304; PC1304.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0019028; C:viral capsid; IEA.		
DR	GO; GO:0019031; C:viral envelope; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR002521; HCV core.		
DR	InterPro; IPR002519; HCV env.		
DR	Pfam; PF01542; HCV_core; 1.		
DR	Pfam; PF01539; HCV_env; 1.		
KW	Coat protein; Envelope protein; Glycoprotein; Polyprotein;		
KW	Transmembrane.		
FT	NON_TER 1		
FT	NON_TER 180		
SQ	SEQUENCE 180 AA; 18934 MW; 15FEE3DE5481693 CRC64;		

Query Match 100.0%; Score 964; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 2.2e-86;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VGAPVGGVARALAHGVRALEDGNGINATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Db	1	VGAPVGGVARALAHGVRALEDGNGINATGNLPGCSFSIFLLALFSCLIHPAASLEWRNTSG 60
Qy	61	LYVLTNDCSNSISVYEADDDVILHTFGCIPCVQDGNSTCTWTPTPTVAKYVGATTASIR 120
Db	61	LYVLTNDCSNSISVYEADDDVILHTFGCIPCVQDGNSTCTWTPTPTVAKYVGATTASIR 120
Qy	121	SHVDLLVGAATWCSALYVGDNGCAFLVYCOAATFFPRRHOTVQTCNCSLYPGHLSGRMA 180
Db	121	SHVDLLVGAATWCSALYVGDNGCAFLVYCOAATFFPRRHOTVQTCNCSLYPGHLSGRMA 180

RESULT 2

Q81752  
ID Q81752 PRELIMINARY; PRT; 503 AA.  
AC Q81752;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCR-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065664; PubMed=7504073;  
RA Okamoto H., Tokita H., Sakamoto M., Horikita M., Kojima M., Iizuka H.,  
RA Mishiro S.;  
RT "Characterization of the genomic sequence of type V (or 3a) hepatitis  
RT C virus isolates and PCR primers for specific detection.";  
RL J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D14305; BAA03251.1; -;  
DR PIR; PQ0804;  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT CHAIN 1 191 C.  
FT CHAIN 192 383 E.  
FT CHAIN 384 >503 E2/NS1.  
FT NON TER 503 503  
SQ SEQUENCE 503 AA; 54822 MW; 24B9CCB7987FB94B CRC64;  
  
Query Match 99.6%; Score 960; DB 2; Length 503;  
Best Local Similarity 98.9%; Pred. No. 1.7e-85;  
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
Db 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
  
Qy 61 LYVLTDNCSSIVYEADVLHTPGICPVQDGNSTCKTPTVTVAVKYVCATTASIR 120  
Db 200 LYVLTDNCSSIVYEADVLHTPGICPVQDGNSTCKTPTVTVAVKYVCATTASIR 259  
  
Qy 121 SHVDLLVGAATMCSALYVGMCGAVFLVQGAFTFRPRRHQTVCNCSLYPGHLSGHRMA 180  
Db 260 SHVDLLVGAATMCSALYVGMCGAVFLVQGAFTFRPRRHQTVCNCSLYPGHLSGHRMA 319  
  
RESULT 3  
Q68870  
ID Q68870 PRELIMINARY; PRT; 3021 AA.  
AC Q68870;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Genes for core, envelope and NS1 proteins.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.

OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Serum;  
RA Seelig R., Weber P., Seeling H.P., Ledger N., Bottner C., Renz M.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X76918; CAAS4244.1; -;  
DR PIR; PC1307; PC1307.  
DR PIR; PQ0401; PQ0401.  
DR PIR; PQ0804; PQ0804.  
DR PIR; S41288; S41288.  
DR HSP; P27958; 1A1V.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005524; F: ATP binding; IEA.  
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F: RNA binding; IEA.  
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P: transcription; IEA.  
DR GO; GO:0019079; P: viral genome replication; IEA.  
DR GO; GO:0019087; P: viral transformation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD\_heme\_BS.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR InterPro; IPR000745; HCV NS4a.  
DR InterPro; IPR001490; HCV NS4b.  
DR InterPro; IPR002568; HCV NS5a.  
DR InterPro; IPR002166; HCV RdRP.  
DR InterPro; IPR004109; Peptidase\_S29.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
DR Pfam; PF01538; HCV NS2; 1.  
DR Pfam; PF02907; HCV NS3; 1.  
DR Pfam; PF01006; HCV NS4a; 1.  
DR Pfam; PF01001; HCV NS4b; 1.  
DR Pfam; PF01506; HCV NS5a; 1.  
DR Pfam; PF00998; Viral\_RdRP; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT CHAIN 736 1012 NS2 protein.  
FT CHAIN 1013 1663 NS3 protein.  
FT CHAIN 1664 1717 NS4a protein.  
FT CHAIN 1 191 core protein.  
FT CHAIN 1718 1978 NS4b protein.  
FT CHAIN 1979 2430 NS5a protein.  
FT CHAIN 2431 3021 NS5b protein.  
FT CHAIN 192 383 envelope protein.  
FT CHAIN 384 735 NS1 protein.  
SQ SEQUENCE 3021 AA; 329093 MW; BF2B499AA5A58CF CRC64;  
  
Query Match 99.6%; Score 960; DB 2; Length 3021;  
Best Local Similarity 98.9%; Pred. No. 1.2e-84;  
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
Db 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199



QY 61 LYLVTNDCSNSSIVYEADVVILHTPGCIPCVQDGNSTCTWPTVPTVAVRYVGATTASIR 120  
 DB 200 LYLVTNDCSNSSIVYEADVVILHTPGCIPCVQDGNSTCTWPTVPTVAVRYVGATTASIR 259  
 QY 121 SHVLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQVTCNCSLYPGHLSGHRMA 180  
 DB 260 SHVLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQVTCNCSLYPGHLSGHRMA 319

RESULT 4  
 Q81258 PRELIMINARY; PRT; 3021 AA.  
 AC Q81258;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sakamoto M.;  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Okamoto H.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D17763; BAA04609.1; -;  
 DR PIR; PQ0401; PQ0401.  
 DR PIR; PQ0804; PQ0804.  
 DR HSSP; P27958; 1HEI.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005234; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR GO; GO:0019079; P:viral genome replication; IEA.  
 DR GO; GO:0019087; P:viral transformation; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR002522; HCV\_capsid.  
 DR InterPro; IPR002521; HCV env.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR InterPro; IPR000745; HCV NS4a.  
 DR InterPro; IPR001490; HCV NS4b.  
 DR InterPro; IPR002868; HCV NS5a.  
 DR InterPro; IPR002166; HCV NS5a.  
 DR InterPro; IPR004109; Peptidase S29.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro; IPR002016; Peroxidase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF01543; HCV\_capsid; 1.  
 DR Pfam; PF01542; HCV\_core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 DR Pfam; PF01538; HCV NS2; 1.  
 DR Pfam; PF02907; HCV NS3; 1.  
 DR Pfam; PF02907; HCV NS3; 1.  
 DR Pfam; PF01006; HCV NS4a; 1.  
 DR Pfam; PF01001; HCV NS4b; 1.  
 DR Pfam; PF01506; HCV NS5a; 1.  
 DR Pfam; PF00998; Viral\_RMRP; 1.  
 DR SMART; SM00487; DEXDC; 1.

DR PROSITE; P800190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE; P800435; PEROXIDASE\_1; UNKNOWN\_1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KM Polyprotein; Transmembrane.  
 FT CHAIN 736 1012 NS2.  
 FT CHAIN 1013 1663 NS3.  
 FT CHAIN 1664 1717 NS4a.  
 FT CHAIN 1 191 C.  
 FT CHAIN 1718 1978 NS4b.  
 FT CHAIN 1979 2430 NS5a.  
 FT CHAIN 2431 3021 NS5b.  
 FT CHAIN 192 383 E1.  
 FT CHAIN 384 735 E2/NS1.  
 SQ SEQUENCE 3021 AA; 329575 MW; 38712CCBC0C19562 CRC64;  
 Query Match 99.6%; Score 960; DB 2; Length 3021;  
 Best Local Similarity 98.9%; Pred. No. 1.2e-84;  
 Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VCAPVGGVARALAHGVRALEDGINFATGNLPCGSFISIFLLALFSLIHPAASLEWRNTSG 60  
 DB 140 VCAPVGGVARALAHGVRALEDGINFATGNLPCGSFISIFLLALFSLIHPAASLEWRNTSG 199  
 QY 61 LYLVTNDCSNSSIVYEADVVILHTPGCIPCVQDGNSTCTWPTVPTVAVRYVGATTASIR 120  
 DB 200 LYLVTNDCSNSSIVYEADVVILHTPGCIPCVQDGNSTCTWPTVPTVAVRYVGATTASIR 259  
 QY 121 SHVLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQVTCNCSLYPGHLSGHRMA 180  
 DB 260 SHVLLVGAATMCSALYVGMCGAVFLVGOAFTFRPRRHQVTCNCSLYPGHLSGHRMA 319

RESULT 5  
 Q81813 PRELIMINARY; PRT; 411 AA.  
 ID Q81813;  
 AC Q81813;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Structural region (Fragment).  
 DE Name-structural region;  
 GN Hepatitis C virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Serum;  
 RX MEDLINE=91172826; PubMed=1848704;  
 RA Choo Q., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,  
 RA Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,  
 RA Bradley D.W., Kuo G., Houghton M.;  
 RT "Genetic organization and diversity of the hepatitis C virus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Serum;  
 RX MEDLINE=91156678; PubMed=1705704;  
 RA Han J.H., Shyamala V., Richman K.H., Brauer M.J., Tekamp-Olson P.;  
 RT "Characterization of the terminal regions of hepatitis C viral RNA:  
 RT Identification of conserved sequences in the 5'-untranslated region  
 RT and.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:1711-1715 (1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Serum;  
 RX MEDLINE=94197744; PubMed=8147893;  
 RA Li J.S., Vitvitski L., Tong S.P., Trepo C.;  
 RT "Identification of the third major genotype of hepatitis C virus in  
 RT France.";  
 RL Biochem. Biophys. Res. Commun. 199:1474-1481 (1994).  
 DR EMBL; L12355; AAA20155.1; -;  
 DR PIR; PC2061; PC2061.

DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; F: structural molecule activity; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro: IPR002522; HCV capsid.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR Pfam: PF01543; HCV capsid; 1.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 Transmembrane.  
 FT NON\_TER 411  
 SQ SEQUENCE 411 AA; 44547 MW; F3C3233E004D6298F CRC64;  
 Query Match 99.1%; Score 955; DB 2; Length 411;  
 Best Local Similarity 97.8%; Pred. No. 4.1e-85;  
 Matches 176; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 Db 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 Db 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 259  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 Db 260 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 6  
 Q81277 PRELIMINARY; PRT; 180 AA.  
 AC Q81277;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core protein-envelope 1 protein (fragment).  
 GN Name=core-E1;  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93249436; PubMed=7683463;  
 RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;  
 RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV  
 type 3";  
 RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
 DR EMBL; D14596; BAA03445.1; -;  
 DR PIR; PC1305; PC1305.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; F: structural molecule activity; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 180  
 SQ SEQUENCE 180 AA; 18887 MW; A24D1A0510CAA9F5 CRC64;  
 Query Match 98.7%; Score 951; DB 2; Length 180;  
 Best Local Similarity 97.8%; Pred. No. 4.2e-85;  
 Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 Db 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 Db 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 259  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 Db 260 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 7  
 Q81279 PRELIMINARY; PRT; 180 AA.  
 AC Q81279;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core protein-envelope 1 protein (fragment).  
 GN Name=core-E1;  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93249436; PubMed=7683463;  
 RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;  
 RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV  
 type 3";  
 RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
 DR EMBL; D14596; BAA03445.1; -;  
 DR PIR; PC1305; PC1305.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; F: structural molecule activity; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 180  
 SQ SEQUENCE 180 AA; 19005 MW; F76270697B0C04F4 CRC64;  
 Query Match 98.3%; Score 948; DB 2; Length 180;  
 Best Local Similarity 97.2%; Pred. No. 8.2e-85;  
 Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 Db 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 RESULT 8  
 Q81489 PRELIMINARY; PRT; 503 AA.  
 AC Q81489;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein (fragment).  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60

Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 Db 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 RESULT 7  
 Q81279 PRELIMINARY; PRT; 180 AA.  
 AC Q81279;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core protein-envelope 1 protein (fragment).  
 GN Name=core-E1;  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93249436; PubMed=7683463;  
 RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;  
 RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV  
 type 3";  
 RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
 DR EMBL; D14596; BAA03445.1; -;  
 DR PIR; PC1305; PC1305.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; F: structural molecule activity; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 180  
 SQ SEQUENCE 180 AA; 19005 MW; F76270697B0C04F4 CRC64;  
 Query Match 98.3%; Score 948; DB 2; Length 180;  
 Best Local Similarity 97.2%; Pred. No. 8.2e-85;  
 Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 Db 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 Db 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYVGATTASIR 120  
 QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 Db 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQTVCNCSLYPGHLSGHRMA 180  
 RESULT 8  
 Q81489 PRELIMINARY; PRT; 503 AA.  
 AC Q81489;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein (fragment).  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60

OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065664; PubMed=7504073;  
RA Okamoto H., Tokita H., Sakamoto M., Horikita M., Kojima M., Iizuka H.,  
RA Mishiro S.;  
RT "Characterization of the genomic sequence of type V (or 3a) hepatitis  
C virus isolates and PCR primers for specific detection.";  
RL J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D14311; BAA03257.1; -;  
DR F1R; PC2060;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002521; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
FT CHAIN 1 191 C.  
FT CHAIN 192 383 E.  
FT CHAIN 384 >503 E2/NS1.  
FT NON TER 503 503  
SQ SEQUENCE 503 AA; 54531 MW; 2FE38EC30F79229 CRC64;  
Query Match 97.9%; Score 944; DB 2; Length 503;  
Best Local Similarity 97.2%; Pred. No. 6.2e-84;  
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
QY 61 LVLTNDCSNSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 200 LVLTNDCSNSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259  
QY 121 SHVDLLVGAATMCSALYVGDGMCAGVFLVGQAFTRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATMCSALYVGDGMCAGVFLVGQAFTRPRRHQTVQTCNSLYPGHLSGHRMA 319  
RESULT 9  
Q81827 PRELIMINARY; PRT; 503 AA.  
AC Q81827  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Polyprotein (fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065664; PubMed=7504073;  
RA Okamoto H., Tokita H., Sakamoto M., Horikita M., Kojima M., Iizuka H.,  
RA Mishiro S.;  
RT "Characterization of the genomic sequence of type V (or 3a) hepatitis

RT C virus isolates and PCR primers for specific detection.";  
RL J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D14309; BAA03255.1; -;  
DR F1R; PQ0804; PQ0804.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
FT CHAIN 1 191 C.  
FT CHAIN 192 383 E.  
FT CHAIN 384 >503 E2/NS1.  
FT NON TER 503 503  
SQ SEQUENCE 503 AA; 54833 MW; 31614329C22BDFEE CRC64;  
Query Match 97.9%; Score 944; DB 2; Length 503;  
Best Local Similarity 97.8%; Pred. No. 6.2e-84;  
Matches 176; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
QY 61 LVLTNDCSNSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 200 LVLTNDCSNSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259  
QY 121 SHVDLLVGAATMCSALYVGDGMCAGVFLVGQAFTRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATMCSALYVGDGMCAGVFLVGQAFTRPRRHQTVQTCNSLYPGHLSGHRMA 319  
RESULT 10  
Q92933 PRELIMINARY; PRT; 3021 AA.  
AC Q92933  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=type 3a;  
RA Shukla D.D., Chaturvedi S., Cao J.Y., Hoynes P.A.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF046866; AAC03058.1; -;  
DR HSP; P27958; 1HEI  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003568; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.

```
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5b.
DR InterPro; IPR004109; Peptidease_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01003; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral_RGRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
SQ SEQUENCE 3021 AA; 328905 MW; D786133B330303CD CRC64;

Query Match 97.7%; Score 942; DB 2; Length 3021;
Best Local Similarity 97.2%; Pred. No. 6.7e-83;
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCHLHPAASLEWNTSG 60
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCHLHPAASLEWNTSG 199

QY 61 LYVLNDCSNSIYVEADVVILHTPGICPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120
DB 200 LYVLNDCSNSIYVEADVVILHTPGICPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDVCGAVFLVGQAFTPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 260 SHVDLLVGAATMCSALYVGDVCGAVFLVGQAFTPRRHQTVQTCNCSLYPGHLSGHRMA 319

RESULT 11
Q81822 PRELIMINARY; PRT; 503 AA.
AC Q81822;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94065664; PubMed=7504073;
RA Okamoto H., Tokita H., Sakamoto M., Horikita M., Kojima M., Iizuka H.,
RA "Characterization of the genomic sequence of type V (or 3a) hepatitis
RT C virus isolates and PCR primers for specific detection.";
RL J. Gen. Virol. 74:2385-2390(1993).
RN [2]

SEQUENCE FROM N.A.
Okamoto H.;
Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
EMBL; D14307; BAA03253.1; -.
PIR; P00804; P00804.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019033; C:viral envelope; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
InterPro; IPR002522; HCV_core.
InterPro; IPR002521; HCV env.
InterPro; IPR002531; HCV NS1.
Pfam; PF01543; HCV capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV NS1; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
CHAIN 1 191
E.
E2/NS1.
NON_TER 384 >503
SEQUENCE 503 AA; 54900 MW; E37704D2A438A828 CRC64;

Query Match 97.0%; Score 935; DB 2; Length 503;
Best Local Similarity 96.1%; Pred. No. 4.7e-83;
Matches 173; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCHLHPAASLEWNTSG 60
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCHLHPAASLEWNTSG 199

QY 61 LYVLNDCSNSIYVEADVVILHTPGICPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120
DB 200 LYVLNDCSNSIYVEADVVILHTPGICPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDVCGAVFLVGQAFTPRRHQTVQTCNCSLYPGHLSGHRMA 180
DB 260 SHVDLLVGAATMCSALYVGDVCGAVFLVGQAFTPRRHQTVQTCNCSLYPGHLSGHRMA 319

RESULT 12
Q81495 PRELIMINARY; PRT; 3021 AA.
AC Q81495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K3a;
MEDLINE=95053917; PubMed=7964640;
RA Yamada N., Manihara K., Mitozaki M., Ohba K., Takada A., Tsutsumi M.,
RA Date T.;
RT "Full-length sequence of the genome of hepatitis C virus type 3a:
RT comparative study with different genotypes.";
RL J. Gen. Virol. 75:3279-3284(1994).
DR EMBL; D28917; BAA06044.1; -.
DR HSP; P27958; HSP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019033; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
```

[2]  
 RN MEDLINE=94186155; PubMed=8138250;  
 RX Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,  
 RA Tanaka T., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in  
 RT patients with chronic liver disease in Nepal.";  
 RL Hepatology 19:805-809(1994).  
 DR EMBL; D16612; BAA04034.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0019028; C: viral capsid; IEA.  
 DR GO; GO:0019031; C: viral envelope; IEA.  
 DR GO; GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV core.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR Pfam; PF01543; HCV capsid; 1.  
 DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON TER 415  
 SQ SEQUENCE 415 AA; 44742 MW; 67ADEB47645BC146 CRC64;  
 Query Match 89.7%; Score 865; DB 2; Length 415;  
 Best Local Similarity 88.3%; Pred. No. 2.8e-76;  
 Matches 159; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LVLTNDCSNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 120  
 DB 200 LVLTNDCSNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 259  
 QY 121 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 GHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 14  
 Q81554 PRELIMINARY; PRT; 415 AA.  
 ID Q81554  
 AC Q81554;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JUN-2003 (TREMBLrel. 01, Last sequence update)  
 DE Core, env. and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94201770; PubMed=8151307;  
 RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,  
 RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Nepal with novel genotypes and their  
 RT classification into the third major group.";  
 RL J. Gen. Virol. 75:931-936(1994).  
 DR EMBL; D16620; BAA04042.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR MEDLINE=94186155; PubMed=8138250;  
 RX Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,  
 RA Tanaka T., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in  
 RT patients with chronic liver disease in Nepal.";  
 RL Hepatology 19:805-809(1994).  
 DR EMBL; D16612; BAA04034.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.

GO; GO:0005198; F: structural molecule activity; IEA.  
 GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
 GO; GO:0006350; P: transcription; IEA.  
 GO; GO:0019079; P: viral genome replication; IEA.  
 GO; GO:0019087; P: viral transformation; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001410; DEAD\_heme\_BS.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV core.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR InterPro; IPR000745; HCV NS4a.  
 DR InterPro; IPR001490; HCV NS4b.  
 DR InterPro; IPR002868; HCV NS5a.  
 DR InterPro; IPR002166; HCV NS5a.  
 DR InterPro; IPR002166; HCV NS5a.  
 DR InterPro; IPR004109; Peptidase\_S29.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF01543; HCV capsid; 1.  
 DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 DR Pfam; PF01538; HCV NS2; 1.  
 DR Pfam; PF02907; HCV NS3; 1.  
 DR Pfam; PF01006; HCV NS4a; 1.  
 DR Pfam; PF01001; HCV NS4b; 1.  
 DR Pfam; PF01506; HCV NS5a; 1.  
 DR Pfam; PF03998; Viral RdRp; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; P800190; CYTOCHROME\_C; UNKNOWN 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON TER 3021  
 SQ SEQUENCE 3021 AA; 328387 MW; A97418FF3C062A4 CRC64;  
 Query Match 96.8%; Score 933; DB 2; Length 3021;  
 Best Local Similarity 96.1%; Pred. No. 5.1e-82;  
 Matches 173; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LVLTNDCSNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 120  
 DB 200 LVLTNDCSNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 259  
 QY 121 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 13  
 Q81546 PRELIMINARY; PRT; 415 AA.  
 ID Q81546  
 AC Q81546;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JUN-2003 (TREMBLrel. 01, Last sequence update)  
 DE Core, env. and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94201770; PubMed=8151307;  
 RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,  
 RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Nepal with novel genotypes and their  
 RT classification into the third major group.";  
 RL J. Gen. Virol. 75:931-936(1994).

[2]  
 RN MEDLINE=94186155; PubMed=8138250;  
 RX Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,  
 RA Tanaka T., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in  
 RT patients with chronic liver disease in Nepal.";  
 RL Hepatology 19:805-809(1994).  
 DR EMBL; D16612; BAA04034.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0019028; C: viral capsid; IEA.  
 DR GO; GO:0019031; C: viral envelope; IEA.  
 DR GO; GO:0005198; F: structural molecule activity; IEA.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV core.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR Pfam; PF01543; HCV capsid; 1.  
 DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON TER 415  
 SQ SEQUENCE 415 AA; 44742 MW; 67ADEB47645BC146 CRC64;  
 Query Match 89.7%; Score 865; DB 2; Length 415;  
 Best Local Similarity 88.3%; Pred. No. 2.8e-76;  
 Matches 159; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LVLTNDSCNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 120  
 DB 200 LVLTNDSCNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 259  
 QY 121 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 GHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 14  
 Q81554 PRELIMINARY; PRT; 415 AA.  
 ID Q81554  
 AC Q81554;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JUN-2003 (TREMBLrel. 01, Last sequence update)  
 DE Core, env. and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94201770; PubMed=8151307;  
 RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,  
 RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Nepal with novel genotypes and their  
 RT classification into the third major group.";  
 RL J. Gen. Virol. 75:931-936(1994).  
 DR EMBL; D16620; BAA04042.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR MEDLINE=94186155; PubMed=8138250;  
 RX Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,  
 RA Tanaka T., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in  
 RT patients with chronic liver disease in Nepal.";  
 RL Hepatology 19:805-809(1994).  
 DR EMBL; D16612; BAA04034.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.

GO; GO:0005198; F: structural molecule activity; IEA.  
 GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
 GO; GO:0006350; P: proteolysis; IEA.  
 GO; GO:0019079; P: viral genome replication; IEA.  
 GO; GO:0019087; P: viral transformation; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001410; DEAD\_heme\_BS.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV core.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR InterPro; IPR000745; HCV NS4a.  
 DR InterPro; IPR001490; HCV NS4b.  
 DR InterPro; IPR002868; HCV NS5a.  
 DR InterPro; IPR002166; HCV NS5a.  
 DR InterPro; IPR004109; Peptidase\_S29.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_psvir.  
 DR Pfam; PF01543; HCV capsid; 1.  
 DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 DR Pfam; PF01538; HCV NS2; 1.  
 DR Pfam; PF02907; HCV NS3; 1.  
 DR Pfam; PF01006; HCV NS4a; 1.  
 DR Pfam; PF01001; HCV NS4b; 1.  
 DR Pfam; PF01506; HCV NS5a; 1.  
 DR Pfam; PF03998; Viral\_RdRp; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 SQ SEQUENCE 3021 AA; 328387 MW; A97418FF3C062A4 CRC64;  
 Query Match 96.8%; Score 933; DB 2; Length 3021;  
 Best Local Similarity 96.1%; Pred. No. 5.1e-82;  
 Matches 173; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 60  
 DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSG 199  
 QY 61 LVLTNDSCNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 120  
 DB 200 LVLTNDSCNSSIVYEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVRYVGATTASIR 259  
 QY 121 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 SHVDLLVGAATWCSALYVGDVCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 319  
 RESULT 13  
 Q81546 PRELIMINARY; PRT; 415 AA.  
 ID Q81546  
 AC Q81546;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JUN-2003 (TREMBLrel. 01, Last sequence update)  
 DE Core, env. and part of E2/NS1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94201770; PubMed=8151307;  
 RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,  
 RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;  
 RT "Hepatitis C virus variants from Nepal with novel genotypes and their  
 RT classification into the third major group.";  
 RL J. Gen. Virol. 75:931-936(1994).

DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; C:viral envelope; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002521; HCV\_core.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_Ns1.  
DR Pfam: PF01543; HCV\_capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_Ns1; 1.  
KW Coat protein, Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.  
FT NON\_TER 415 415  
SQ SEQUENCE 415 AA; 44839 MW; 809114E3D485C397 CRC64;  
  
Query Match 89.1%; Score 859; DB 2; Length 415;  
Best Local Similarity 87.8%; Pred. No. 1.1e-75;  
Matches 158; Conservative 9; Mismatches 13; Indels 0; Gaps 0;  
  
QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
  
QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYGATTASIR 120  
DB 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYGATTASIR 259  
  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTPRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATLCSALYVGDGCGAVFLVGOAFTPRPRRHQTVQTCNCSLYPGHLSGHRMA 319  
  
RESULT 15  
Q81552  
ID Q81552 PRELIMINARY; PRT; 415 AA.  
AC Q81552;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Core, env, and part of E2/Ns1 (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94201770; PubMed=8151307;  
RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,  
RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;  
RT "Hepatitis C virus variants from Nepal with novel genotypes and their  
RT classification into the third major group.";  
RL J. Gen. Virol. 75:931-936(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94186155; PubMed=8138250;  
RA Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,  
RA Tanaka T., Miyakawa Y., Mayumi M.;  
RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in  
RT patients with chronic liver disease in Nepal.";  
RL Hepatology 19:805-809(1994).  
DR EMBL: D16618; BAA0404.1; -;  
DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; C:viral envelope; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002521; HCV\_core.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_Ns1.  
DR Pfam: PF01543; HCV\_Ns1.  
DR Pfam: PF01542; HCV\_capsid; 1.  
DR Pfam: PF01543; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.

DR Pfam: PF01560; HCV\_Ns1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.  
FT NON\_TER 415 415  
SQ SEQUENCE 415 AA; 44580 MW; 35F4EF4DF6F79C30 CRC64;  
  
Query Match 88.9%; Score 857; DB 2; Length 415;  
Best Local Similarity 86.7%; Pred. No. 1.7e-75;  
Matches 156; Conservative 12; Mismatches 12; Indels 0; Gaps 0;  
  
QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
  
QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYGATTASIR 120  
DB 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGTSTCTWPTPTVAVKYGATTASIR 259  
  
QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTPRPRRHQTVQTCNCSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATLCSALYVGDGCGAVFLVGOAFTPRPRRHQTVQTCNCSLYPGHLSGHRMA 319  
  
Search completed: December 1, 2004, 23:44:13  
Job time : 195 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:35:24 ; Search time 38 Seconds  
(without alignments)  
455.764 Million cell updates/sec

Title: US-09-899-046a-20

Perfect score: 964  
Sequence: 1 VGAPVGVARALAHGVRALE.....TVQTCNCSLYPGHLSGHRMA 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79:\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	964	100.0	180	PC1304	genome polyprotein
2	960	99.6	492	S41288	genome polyprotein
3	955	99.1	411	PC2061	genome polyprotein
4	951	98.7	180	PC1305	genome polyprotein
5	948	98.3	180	PC1303	genome polyprotein
6	923	95.7	411	PC2060	genome polyprotein
7	733	76.0	513	PC1284	genome polyprotein
8	731	75.8	782	S18032	genome polyprotein
9	728	75.5	782	S19876	genome polyprotein
10	727	75.4	3010	S18030	genome polyprotein
11	724	75.1	3010	1 GNVVTC	genome polyprotein
12	723	75.0	3011	1 GNVVCH	genome polyprotein
13	722	74.9	441	S12707	genome polyprotein
14	722	74.9	3010	1 GNVVTV	genome polyprotein
15	721	74.8	782	S18031	genome polyprotein
16	721	74.8	3010	1 GNVVCJ	genome polyprotein
17	717	74.4	513	A44150	structural protein
18	717	74.4	787	2 PNO677	hypothetical prote
19	717	74.4	3011	1 GNVVC3	genome polyprotein
20	717	74.4	3011	1 S40770	genome polyprotein
21	715	74.2	3010	1 A45573	genome polyprotein
22	714	74.1	550	2 JH0711	genome polyprotein
23	713	74.0	520	2 JH1925	polyprotein - hepa
24	712	73.9	369	2 S21471	genome polyprotein
25	711	73.8	782	2 S19875	genome polyprotein
26	708	73.4	523	2 JQ1926	polyprotein - hepa
27	703	72.9	640	2 JQ1584	genome polyprotein
28	666	69.1	3014	1 JCS620	genome polyprotein
29	664	68.9	322	2 JN0265	genome polyprotein

## RESULT 1

PC1304  
genome polyprotein core/E1 region (isolate BR36-9) - hepatitis C virus (strain type 3) (1  
C;Species: hepatitis C virus  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: PC1304  
R;Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.  
Biochem. Biophys. Res. Commun. 192, 635-641, 1993  
A;Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.  
A;Reference number: PC1300; MUID:93249436; PMID:7683463  
A;Accession: PC1304  
A;Molecule type: mRNA  
A;Residues: 1-180 <STU>  
A;Cross-references: UNIPROT:Q81278; DBJ:D14599; NID:G303578; PIDN:BAA03448.1; PID:G303578  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p  
F;53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status Predicted  
F;125-154/Domain: transmembrane #status predicted <TM>  
F;57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 964; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.7e-81;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAAASLEWNTSG	60
Db	1	VGAPVGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAAASLEWNTSG	60
Qy	61	LYVLTNDCSNSISIVYEADVVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR	120
Db	61	LYVLTNDCSNSISIVYEADVVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTASIR	120
Qy	121	SHVDLLVGATWCALYVGMCGAVFLVGQAFERRRHHQTQTCNCSLYPGHLSGHRVA	180
Db	121	SHVDLLVGATWCALYVGMCGAVFLVGQAFERRRHHQTQTCNCSLYPGHLSGHRVA	180

## RESULT 2

S41288

Genome polyprotein - hepatitis C virus (fragment)

N;Contains: core protein; envelope protein; NS1 protein

C;Species: hepatitis C virus

C;Date: 06-Jan-1995 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004

C;Accession: S41288

R;Seelig, R.

submitted to the EMBL Data Library, December 1993

A;Reference number: S41288

A;Accession: S41288

A;Molecule type: genomic RNA

A;Residues: 1-492 <SEE>

A;Cross-references: UNIPROT:Q68870; EMBL:X76918

C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; core protein; envelope protein; nonstructural protein; polyprotein  
F:1191/Product: core protein #status predicted <COR>  
F:192-372/Product: envelope protein #status predicted <ENV>  
F:373-492/Product: NS1 protein (fragment) #status predicted <NS1>

Query Match 99.6%; Score 960; DB 2; Length 492;  
Best Local Similarity 98.9%; Pred. No. 1.1e-80;  
Matches 178; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 199

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 319

RESULT 3  
PC2061  
genome polyprotein N2 - hepatitis C virus  
N:Contains: envelope protein E1; nonstructural protein E2/NS1  
C:Species: hepatitis C virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: PC2061  
R:Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.  
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994  
A:Title: Identification of the third major genotype of hepatitis C virus in France.  
A:Reference number: PC2060; MUID:94197744; PMID:8147893  
A:Accession: PC2061  
A:Molecule type: mRNA  
A:Residues: 1-411 <L1J>  
A:Cross-references: UNIPROT:Q81813; GB:L12355; NID:G410169; PIDN:AAA20155.1; PID:G410170  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:192-383/Product: envelope protein E1 #status predicted <SPB>  
F:384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>  
F:196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.1%; Score 955; DB 2; Length 411;  
Best Local Similarity 97.8%; Pred. No. 2.6e-80;  
Matches 176; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 199

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 319

RESULT 4  
PC1305  
genome polyprotein core/E1 region (isolate BR33-1) - hepatitis C virus (strain type 3)  
C:Species: hepatitis C virus  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C:Accession: PC1305  
R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; DeLeys, R.; Maertens, G.  
Biochem. Biophys. Res. Commun. 192, 635-641, 1993  
A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.  
A:Reference number: PC1300; MUID:93249436; PMID:7683463  
A:Accession: PC1305  
A:Molecule type: mRNA

A:Residues: 1-180 <STU>  
A:Cross-references: UNIPROT:Q81277; DDBJ:D14596; NID:G303576; PIDN:BAA03445.1; PID:G30358  
A:Experimental source: blood  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p;  
F:53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status predicted  
F:125-154/Domain: transmembrane #status predicted <TMM>  
F:57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.7%; Score 951; DB 2; Length 180;  
Best Local Similarity 97.8%; Pred. No. 2.6e-80;  
Matches 176; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180

RESULT 5  
PC1303  
genome polyprotein core/E1 region (isolate HD10-2) - hepatitis C virus (strain type 3) (i  
C:Species: hepatitis C virus  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C:Accession: PC1303  
R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; DeLeys, R.; Maertens, G.  
Biochem. Biophys. Res. Commun. 192, 635-641, 1993  
A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.  
A:Reference number: PC1300; MUID:93249436; PMID:7683463  
A:Accession: PC1303  
A:Molecule type: mRNA  
A:Residues: 1-180 <STU>  
A:Cross-references: UNIPROT:Q81279; DDBJ:D14603; NID:G303580; PIDN:BAA03452.1; PID:G30358  
A:Experimental source: blood  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane p;  
F:53-64,74-84,91-103,109-118,155-164/Region: hypervariable region #status predicted  
F:125-154/Domain: transmembrane #status predicted <TMM>  
F:57,70,95,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.3%; Score 948; DB 2; Length 180;  
Best Local Similarity 97.2%; Pred. No. 4.9e-80;  
Matches 175; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60  
DB 1 VGAPVGGVARALAHGVRALEDGINFATGNLPFGCSFSIFLLALFSLCLHPAASLEWRNTSG 60

QY 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120  
DB 61 LYVLTNDCSNSIYVEADDVILHTPGCIPCVQDGNSTCTWPTPTVAVKYVGATTASIR 120

QY 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 121 SHVDLLVGAATMCSALYVGMCGAVFLVGQAFTFRPRRHQTVQTCNSLYPGHLSGHRMA 180

RESULT 6  
PC2060  
genome polyprotein N1 - hepatitis C virus  
N:Contains: envelope protein E1; nonstructural protein E2/NS1  
C:Species: hepatitis C virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: PC2060  
R:Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.  
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994



A>Title: Identification of the third major genotype of hepatitis C virus in France.  
A/Reference number: PC2060; MUID:94197744; PMID:8147893  
A/Accession: PC2060  
A/Molecule type: mRNA  
A/Residues: 1-411 <LIU>  
A/Cross-references: UNIPROT:Q81489  
A/Superfamily: hepatitis C virus genome polyprotein  
C/Keywords: AT; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F;192-383/Product: envelope protein E1 #status predicted <SPS>  
F;384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>  
F;196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.7%; Score 923; DB 2; Length 411;  
Best Local Similarity 95.6%; Pred. No. 2.3e-77;  
Matches 172; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 180  
DB 260 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 319

RESULT 7  
PC1284  
genome polyprotein - hepatitis C virus (isolate HC-J4) (fragment)  
C/Species: hepatitis C virus  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
A/Accession: PC1284  
R/Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A/Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A/Reference number: PC1284; MUID:91013116; PMID:2170712  
A/Accession: PC1284  
A/Molecule type: genomic RNA  
A/Residues: 1-513 <OKA>  
A/Cross-references: UNIPROT:Q81221; GB:D00832; NID:g221513; PIDN:BAA00706.1; PID:g221514  
A/Superfamily: hepatitis C virus genome polyprotein

Query Match 76.0%; Score 733; DB 2; Length 513;  
Best Local Similarity 72.2%; Pred. No. 8.8e-60;  
Matches 130; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 180  
DB 260 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 319

RESULT 8  
S18032  
genome polyprotein - hepatitis C virus (isolate JK4) (fragment)  
N/Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C/Species: hepatitis C virus  
A/Variety: isolate JK4  
C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
A/Accession: S18032  
R/Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991  
A/Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso

A/Reference number: S18029  
A/Accession: S18032  
A/Molecule type: genomic RNA  
A/Residues: 1-782 <HON>  
A/Cross-references: UNIPROT:Q68952; EMBL:X61594  
A/Experimental source: isolate JK4  
A/Superfamily: hepatitis C virus genome polyprotein  
C/Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F;1-191/Product: core protein #status predicted <MAT1>  
F;192-383/Product: envelope protein 1 #status predicted <MAT2>  
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 75.8%; Score 731; DB 2; Length 782;  
Best Local Similarity 72.2%; Pred. No. 2.1e-59;  
Matches 130; Conservative 24; Mismatches 26; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 180  
DB 260 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 319

RESULT 9  
S19876  
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)  
N/Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C/Species: hepatitis C virus  
A/Variety: isolate JK5  
C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
A/Accession: S19876  
R/Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991  
A/Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso

A/Reference number: S18029  
A/Accession: S19876  
A/Molecule type: genomic RNA  
A/Residues: 1-782 <HON>  
A/Cross-references: UNIPROT:Q68953; EMBL:X61595; NID:g59486; PIDN:CAA43792.1; PID:g59487  
A/Experimental source: isolate JK5  
A/Superfamily: hepatitis C virus genome polyprotein  
C/Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F;1-191/Product: core protein #status predicted <MAT1>  
F;192-383/Product: envelope protein 1 #status predicted <MAT2>  
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 75.5%; Score 728; DB 2; Length 782;  
Best Local Similarity 72.8%; Pred. No. 3.9e-59;  
Matches 131; Conservative 21; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPVGGVARALAHGVRVALEGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 120  
DB 200 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCWTPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 180  
DB 260 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQVOTCNCSLYPGHLGSHRMA 319

RESULT 10  
S18030

genome polyprotein - hepatitis C virus (isolate JKI)  
 N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein  
 Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: hepatitis C virus  
 A:Variety: isolate JKI  
 C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
 A:Accession: S18030; S33570; S18029  
 R:Honda, M.; Kaneko, S.; Magashi, U.; Kobayashi, K.; Murakami, S.  
 submitted to the EMBL Data Library, September 1991  
 A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient  
 A:Reference number: S18028  
 A:Accession: S18030  
 A:Molecule type: genomic RNA  
 A:Residues: 1-3010 <HON>  
 A:Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:G59478; PIDN:CAA43793.1; PID:G59478  
 A:Experimental sources: isolate JKI from an individual  
 R:Honda, M.; Kaneko, S.; Uchida, M.; Kobayashi, K.; Murakami, S.  
 Arch. Virol. 128, 163-169, 1993  
 A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated  
 A:Reference number: A48332; MUID:93119270; PMID:8380322  
 A:Accession: S33570  
 A:Molecule type: genomic RNA  
 A:Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>  
 A:Cross-references: EMBL:X61591  
 A:Note: this sequence is inconsistent with the nucleotide translation  
 A:Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320  
 as Trp, and TTC for residue 771 as Ser  
 A:Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:121748)  
 C:Superfamily: hepatitis C virus genome polyprotein  
 C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin  
 F:115/Product: capsid protein C #status predicted <CPC>  
 F:116-191/Product: envelope protein M #status predicted <EPM>  
 F:192-389/Product: major envelope protein E #status predicted <MEE>  
 F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
 F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>  
 F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
 F:1312-1317/Region: nucleotide-binding motif B  
 F:1316-1319/Region: DEXH motif  
 F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
 F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
 F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:196, 209, 234, 250, 305, 417, 423, 448, 532, 540, 556, 576, 623, 645/Binding site: carbohydrate (As  
 Query Match 75.4%; Score 727; DB 1; Length 3010;  
 Best Local Similarity 71.7%; Pred. No. 1.9e-58;  
 Matches 129; Conservative 24; Mismatches 27; Indels 0; Gaps 0;  
 QY 1 VGAPVGVARALAHGVRALEDGINFATNPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 140 VGAPLGGARALAHGVRALEDGVNATGNLPGCSFSIFLLALLSCLTTPVSTVEVRVNSG 199  
 QY 61 LYVLTNDCNSISIVYEAADVILHTPGICPCVQDGNSTCTVPTVAVKYVGATTASIR 120  
 DB 200 VYHTVNDNSNSISIVYEAADIMHTPGVPCVREGNSRCWALPTTLAARNSIPTTTIR 259  
 QY 121 SHVDLLVGAATWCSALYVGMCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 RHVDLLVGAALCSAMVVDLCSGVFLVSQFLTFSRPRYRTVQDNCNLSYFGHVSGRMA 319  
 RESULT 11  
 GNNVTC  
 Genome polyprotein - hepatitis C virus  
 N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein  
 Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: hepatitis C virus  
 A:Variety: isolate JKI  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
 A:Accession: A38465  
 R:Takamizawa, A.; Moxi, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;  
 J. Virol. 65, 1105-1113, 1991  
 A:Title: Structure and organization of the hepatitis C virus genome isolated from human  
 A:Reference number: A38465; MUID:91140698; PMID:1847440

A:Accession: A38465  
 A:Molecule type: genomic RNA  
 A:Residues: 1-3010 <TK>  
 A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:G329770; PIDN:AAA72945.1; PID:G32977;  
 C:Superfamily: hepatitis C virus genome polyprotein  
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
 F:115/Product: capsid protein C #status predicted <CPC>  
 F:116-191/Product: envelope protein M #status predicted <EPM>  
 F:192-389/Product: major envelope protein E #status predicted <MEE>  
 F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
 F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>  
 F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
 F:1312-1317/Region: nucleotide-binding motif B  
 F:1316-1319/Region: DEXH motif  
 F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
 F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
 F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:196, 209, 234, 250, 305, 417, 423, 448, 532, 540, 556, 576, 623, 645/Binding site: carbohydrate (As  
 Query Match 75.1%; Score 724; DB 1; Length 3010;  
 Best Local Similarity 71.1%; Pred. No. 3.6e-58;  
 Matches 128; Conservative 24; Mismatches 28; Indels 0; Gaps 0;  
 QY 1 VGAPVGVARALAHGVRALEDGINFATNPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
 DB 140 VGAPLGGARALAHGVRALEDGVNATGNLPGCSFSIFLLALLSCLTTPASAYEVHNSG 199  
 QY 61 LYVLTNDCNSISIVYEAADVILHTPGICPCVQDGNSTCTVPTVAVKYVGATTASIR 120  
 DB 200 VYHTVNDNSNSISIVYEAADIMHTPGVPCVREGNSRCWALPTTLAARNSIPTTTIR 259  
 QY 121 SHVDLLVGAATWCSALYVGMCGAVFLVGOAFTPRPRHQTVCNCSLYPGHLSGHRMA 180  
 DB 260 RHVDLLVGAALCSAMVVDLCSGVFLVSQFLTFSRPRHVTLDQNCNLSYFGHVSGRMA 319  
 RESULT 12  
 GNNVTC  
 Genome polyprotein - hepatitis C virus (strain H)  
 N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein  
 Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: hepatitis C virus  
 A:Variety: host Homo sapiens (man)  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 A:Accession: A36814; A41546  
 R:Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.  
 submitted to GenBank, July 1992  
 A:Description: Genomic structure of the human prototype strain H of hepatitis C virus: cc  
 A:Reference number: A36814  
 A:Accession: A36814  
 A:Molecule type: genomic RNA  
 A:Residues: 1-3011 <INC>  
 A:Cross-references: UNIPROT:P27958; GB:M67463; NID:G329737; PIDN:AAA45534.1; PID:G329738  
 R:Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991  
 A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari  
 A:Reference number: A41546; MUID:92052256; PMID:1658800  
 A:Contents: annotation  
 A:Note: neither amino acid nor nucleotide sequence is given  
 C:Superfamily: hepatitis C virus genome polyprotein  
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
 F:115/Product: capsid protein C #status predicted <CPC>  
 F:116-191/Product: envelope protein M #status predicted <EPM>  
 F:192-389/Product: major envelope protein E #status predicted <MEE>  
 F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
 F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>  
 F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
 F:1312-1317/Region: nucleotide-binding motif B  
 F:1316-1319/Region: DEXH motif  
 F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
 F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match 75.0%; Score 723; DB 1; Length 3011;  
Best Local Similarity 71.1%; Pred. No. 4.5e-58;  
Matches 128; Conservative 22; Mismatches 30; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPLGGARALAHGVRVLEDDGVNATGNLPGCSFSIFLLALLSCLTIPASAYEVNRSSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 120  
DB 200 LYHVTNDCNSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 RHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 319

RESULT 13  
S12707  
genome polyprotein - hepatitis C virus (fragment)  
N:Contains: core protein; envelope protein  
C:Species: hepatitis C virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: S12707  
R:Takeuchi, K.; Kubo, Y.; Boonmar, S.; Watanabe, Y.; Katayama, T.; Choo, Q.L.; Kuo, G.;  
Nucleic Acids Res. 18, 4626, 1990  
A:Title: Nucleotide sequence of core and envelope genes of the hepatitis C virus genome  
A:Reference number: S12707; UID:90356432; PMID:2117749  
A:Accession: S12707  
A:Molecule type: genomic RNA  
A:Residues: 1-441 <TAK>  
A:Cross-references: UNIPROT:Q81776; EMBL:D00574; NID:G221656; PIDN:BAA00452.1; PID:G2216  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: polyprotein

Query Match 74.9%; Score 722; DB 2; Length 441;  
Best Local Similarity 70.0%; Pred. No. 7.8e-59;  
Matches 126; Conservative 24; Mismatches 30; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPLGGARALAHGVRVLEDDGVNATGNLPGCSFSIFLLALLSCLTIPASAYEVNRSSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 120  
DB 200 LYHVTNDCNSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 RHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 319

RESULT 14  
GNWVTW  
genome polyprotein - hepatitis C virus (strain Taiwan)  
N:Contains: capsid protein C; envelope protein E; hepatitis C virus genome; nonstructural  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: A40244  
R:Chen, P. J.; Lin, M. H.; Tai, K. F.; Liu, P. C.; Lin, C. J.; Chen, D. S.  
Virology 188, 102-113, 1992  
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the  
A:Reference number: A40244; UID:92230206; PMID:1314449  
A:Accession: A40244  
A:Molecule type: genomic RNA  
A:Residues: 1-3010 <CHE>  
A:Cross-references: UNIPROT:P29846; GB:M84754  
C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
F:1-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MES>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,207

Query Match 74.9%; Score 722; DB 1; Length 3010;  
Best Local Similarity 71.1%; Pred. No. 5.5e-58;  
Matches 128; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPLGGARALAHGVRVLEDDGVNATGNLPGCSFSIFLLALLSCLTIPASAYEVNRSSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 120  
DB 200 LYHVTNDCNSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 RHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 319

RESULT 15  
S18031  
genome polyprotein - hepatitis C virus (isolate JK2) (fragment)  
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C:Species: hepatitis C virus  
A:Variety: isolate JK2  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: S18031  
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.  
Submitted to the EMBL Data Library, September 1991  
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso:  
A:Reference number: S18029  
A:Accession: S18031  
A:Molecule type: Genomic RNA  
A:Residues: 1-782 <HON>  
A:Cross-references: UNIPROT:Q68950; EMBL:X61593  
A:Experimental source: isolate JK2  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F:1-191/Product: core protein #status predicted <M1>  
F:192-383/Product: envelope protein 1 #status predicted <M2>  
F:384-733/Product: NS1/E2 protein #status predicted <M3>  
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <M4>

Query Match 74.8%; Score 721; DB 2; Length 782;  
Best Local Similarity 72.2%; Pred. No. 1.7e-58;  
Matches 130; Conservative 21; Mismatches 29; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRVALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
DB 140 VGAPLGGARALAHGVRVLEDDGVNATGNLPGCSFSIFLLALLSCLTIPASAYEVNRSSG 199

QY 61 LYVLTNDCNSSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 120  
DB 200 LYHVTNDCNSIYVEADVVILHTPGICPCVODGNTSTCTWPTVPTVAVKYVGATTASIR 259

QY 121 SHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 180  
DB 260 RHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFPRRHQTVQTCNSLYPGHLSGHRMA 319

Thu Dec 2 10:55:53 2004

us-09-899-046a-20.rpr

Page 6

Search completed: December 1, 2004, 23:44:56  
Job time : 39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2004, 23:44:21 ; Search time 193 Seconds  
(without alignments)

536.619 Million cell updates/sec

Title: US-09-899-046A-20

Perfect score: 180

Sequence: 1 VGAPVGVARALAHGVRALE.....TVQTCNCSLYPGHLSGHMA 180

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	100.0	180	2	Q81278
2	101	56.1	3021	2	O81495
3	91	50.6	117	2	O80KX0
4	87	48.3	180	2	O81279
5	87	48.3	503	2	O81489
6	87	48.3	503	2	O81252
7	87	48.3	3021	2	O92933
8	87	48.3	3021	2	O68870
9	87	48.3	3021	2	O81258
10	85	47.2	149	2	O68728
11	78	43.3	128	2	O68176
12	78	43.3	128	2	O68178
13	70	38.9	119	2	O80NE2
14	70	38.9	128	2	O68185
15	70	38.9	180	2	O81277
16	70	38.9	192	2	O81385
17	70	38.9	192	2	O81423
18	70	38.9	192	2	O9E8X8
19	70	38.9	503	2	O81827
20	68	37.8	117	2	O80KX1
21	68	37.8	119	2	O80NE3
22	67	37.2	149	2	O68738
23	66	36.7	117	2	O80KX3
24	66	36.7	411	2	O81813
25	64	35.6	119	2	O80NE5
26	64	35.6	128	2	O68179
27	64	35.6	128	2	O68180
28	64	35.6	128	2	O68181
29	64	35.6	128	2	O68187
30	63	35.0	128	2	O68183
31	62	34.4	117	2	O80KX9

32 62 34.4 503 2 Q81822 hepatitis c  
33 60 33.3 117 2 Q80KQ3 hepatitis c  
34 60 33.3 117 2 Q80KQ6 hepatitis c  
35 60 33.3 117 2 Q80KR8 hepatitis c  
36 60 33.3 117 2 Q80KS2 hepatitis c  
37 60 33.3 117 2 Q80KS9 hepatitis c  
38 58 32.2 117 2 Q80KQ4 hepatitis c  
39 58 32.2 117 2 Q80KX5 hepatitis c  
40 53 29.4 117 2 Q80KR6 hepatitis c  
41 52 28.9 191 2 Q6QYU5 hepatitis c  
42 52 28.9 191 2 Q6QYU6 hepatitis c  
43 52 28.9 191 2 Q6QYU7 hepatitis c  
44 52 28.9 191 2 Q6QYU8 hepatitis c  
45 52 28.9 191 2 Q6QYU9 hepatitis c

#### ALIGNMENTS

##### RESULT 1

Q81278 PRELIMINARY; PRT; 180 AA.  
AC Q81278;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Core protein-envelope 1 protein (Fragment).  
GN Name-core-E1;  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92249436; PubMed=7683463;  
RA Stuyver L., Arnhem W.V., Wyseur A., Delays R., Maertens G.;  
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV  
type 3.";  
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
DR EMBL; D14599; BAA03448.1; -;  
DR PIR; PC1304; PC1304  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV\_env.  
DR Pfam; PF01542; HCV\_core\_1.  
DR Pfam; PF01539; HCV\_env; 1.  
DR Coat protein; Envelope protein; Glycoprotein; Polypeptide;  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 180  
SQ SEQUENCE 180 AA; 18934 MW; 15FEE3DE5481693 CRC64;

Query Match 100.0%; Score 180; DB 2; Length 180;

Best Local Similarity 100.0%; Pred. No. 4.7e-172;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGVARALAHGVRALEGGINPFCGSPSIFLLALFSLIHPAASLEWRNTSG 60  
DB 1 VGAPVGVARALAHGVRALEGGINPFCGSPSIFLLALFSLIHPAASLEWRNTSG 60  
QY 61 LYVLNDGNSISVTEADVVILHTPGCIPCVQDGNSTCTWPTVAVKVGATTASIR 120  
DB 61 LYVLNDGNSISVTEADVVILHTPGCIPCVQDGNSTCTWPTVAVKVGATTASIR 120  
QY 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHMA 180  
DB 121 SHVDLLVGAATWCSALYVGDGMCVAVFLVGOAFTFRPRRHQTVQTCNCSLYPGHLSGHMA 180

##### RESULT 2

Q81495 PRELIMINARY; PRT; 3021 AA.  
 ID Q81495  
 AC Q81495  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OC NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K3a;  
 RA Yamada N., Manihara K., Mizokami M., Ohba K., Takada A., Tsutsuami M.,  
 RA Date T.,  
 RA "Full-length sequence of the genome of hepatitis C virus type 3a:  
 RT comparative study with different genotypes.";  
 RL J. Gen. Virol. 75:3279-3284(1994).  
 DR EMBL; D28917; BAA06044.1; -.  
 DR HSP; P27958; IHEI.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR GO; GO:0019079; P:viral genome replication; IEA.  
 DR GO; GO:0019087; P:viral genome replication; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR002522; HCV capsid.  
 DR InterPro; IPR002521; HCV core.  
 DR InterPro; IPR002519; HCV env.  
 DR InterPro; IPR002531; HCV NS1.  
 DR InterPro; IPR000745; HCV NS4a.  
 DR InterPro; IPR001490; HCV NS4b.  
 DR InterPro; IPR002868; HCV NS5a.  
 DR InterPro; IPR002166; HCV RdRp.  
 DR InterPro; IPR004109; Peptidase\_S29.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF01543; HCV capsid; 1.  
 DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Pfam; PF01560; HCV NS1; 1.  
 DR Pfam; PF01538; HCV NS2; 1.  
 DR Pfam; PF02907; HCV NS3; 1.  
 DR Pfam; PF01006; HCV NS4a; 1.  
 DR Pfam; PF01001; HCV NS4b; 1.  
 DR Pfam; PF01506; HCV NS5a; 1.  
 DR Pfam; PF00999; Viral\_RdRp; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
 DR Coated protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 SQ SEQUENCE 3021 AA; 328387 MW; A97418FF36C062A4 CRC64;  
 Query Match 56.1%; Score 101; DB 2; Length 3021;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-92;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 ARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCIHPAASLEWRNTSGLYVLTND 68  
 DB 148 ARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLCIHPAASLEWRNTSGLYVLTND 207

QY 69 SNSSIVYEADVVILHTPGCIPCVODGNTSTCWTPTVPTAV 109  
 DB 208 SNSSIVYEADVVILHTPGCIPCVODGNTSTCWTPTVPTAV 248  
 RESULT 3  
 Q80KSO PRELIMINARY; PRT; 117 AA.  
 ID Q80KSO  
 AC Q80KSO  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Polyprotein (Fragment).  
 OS Hepatitis C virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OC NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22511041; PubMed=12624035;  
 RA Corbet S., Bukh J., Heinsen A., Fomsgaard A.;  
 RT "Hepatitis C virus subtyping by a core-envelope 1-based reverse  
 RT transcriptase PCR assay with sequencing and its use in determining  
 RT subtype distribution among Danish patients.";  
 RL J. Clin. Microbiol. 41:1091-1100(2003).  
 DR EMBL; AY177863; AA016759.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002519; HCV env.  
 DR Pfam; PF01539; HCV env; 1.  
 DR Coated protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1 117  
 FT NON\_TER 117 117  
 SQ SEQUENCE 117 AA; 12652 MW; 7329E0E12EEB4D20 CRC64;  
 Query Match 50.6%; Score 91; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-83;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 80 VILHTPGCIPCVODGNTSTCWTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVG 139  
 DB 27 VILHTPGCIPCVODGNTSTCWTPTVAVKYVGATTASIRSHVDLLVGAATMCSALYVG 86  
 RESULT 4  
 Q81279 PRELIMINARY; PRT; 180 AA.  
 ID Q81279  
 AC Q81279  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Core protein-envelope 1 protein (Fragment).  
 GN Name=core-E1;  
 OS Hepatitis C virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OC NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=93249436; PubMed=7683463;  
 RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;  
 RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV  
 RT type 3.";  
 RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
 DR EMBL; D14603; BAA03452.1; -.

DR PIR; PC1303; PC1303.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polypeptide;  
KW Transmembrane.  
FT NON\_TER 180 180  
SQ SEQUENCE 180 AA; 19005 MW; F76270697B0C04F4 CRC64;  
  
Query Match 48.3%; Score 87; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.1e-78;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
Db 1 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
  
Qy 61 LVYLTNDCSNSIYVEADVVILHTPGC 87  
Db 61 LVYLTNDCSNSIYVEADVVILHTPGC 87  
  
RESULT 5  
Q81489 PRELIMINARY; PRT; 503 AA.  
AC Q81489;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OC NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065664; PubMed=7504073;  
RA Okamoto H.; Tokita H.; Sakamoto M.; Horikita M.; Kojima M.; Iizuka H.;  
RA Mishiro S.;  
RT "Characterization of the genomic sequence of type V (or 3a) hepatitis  
C virus isolates and PCR primers for specific detection.";  
RT J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
RL EMBL; D14311; BAA03257.1; -.  
DR PIR; PC2060; PC2060.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
RL EMBL; D14311; BAA03257.1; -.  
DR PIR; PC2060; PC2060.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polypeptide; Transmembrane.  
FT CHAIN 1 191  
FT CHAIN 192 383  
FT CHAIN 384 >503  
FT NON\_TER 503 503  
SQ SEQUENCE 503 AA; 54531 MW; 2FE38EC30F779229 CRC64;

Query Match 48.3%; Score 87; DB 2; Length 503;  
Best Local Similarity 100.0%; Pred. No. 2.5e-78;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
Db 140 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
  
Qy 61 LVYLTNDCSNSIYVEADVVILHTPGC 87  
Db 200 LVYLTNDCSNSIYVEADVVILHTPGC 226  
  
RESULT 6  
Q81752 PRELIMINARY; PRT; 503 AA.  
AC Q81752;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OC NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065664; PubMed=7504073;  
RA Okamoto H.; Tokita H.; Sakamoto M.; Horikita M.; Kojima M.; Iizuka H.;  
RA Mishiro S.;  
RT "Characterization of the genomic sequence of type V (or 3a) hepatitis  
C virus isolates and PCR primers for specific detection.";  
RT J. Gen. Virol. 74:2385-2390(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Okamoto H.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
RL EMBL; D14305; BAA03251.1; -.  
DR PIR; PQ0804; PQ0804.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0019028; C: viral capsid; IEA.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0005198; F: structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV capsid.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polypeptide; Transmembrane.  
FT CHAIN 1 191  
FT CHAIN 192 383  
FT CHAIN 384 >503  
FT NON\_TER 503 503  
SQ SEQUENCE 503 AA; 54822 MW; 24B9CCB7987FB94B CRC64;  
  
Query Match 48.3%; Score 87; DB 2; Length 503;  
Best Local Similarity 100.0%; Pred. No. 2.5e-78;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60  
Db 140 VGAPVGGVARALAHGVALEEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199  
  
Qy 61 LVYLTNDCSNSIYVEADVVILHTPGC 87  
Db 200 LVYLTNDCSNSIYVEADVVILHTPGC 226

```

RESULT 7
ID 092933 PRELIMINARY; PRT: 3021 AA.
AC 092933;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RC STRAIN-type 3a;
RA Shukla D.D., Chaturvedi S., Cao J.Y., Hoyne P.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF046866; AAC03058.1; -.
DR HSSP; P27958; 1HEI.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0005524; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0008026; F: RNA binding; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3021 AA; 328905 MW; D7B6133B330303CD CRC64;

Query Match 48.3%; Score 87; DB 2; Length 3021;
Best Local Similarity 100.0%; Pred. No. 1.1e-77;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSLFLALFSCULHPAASLEWNTSG 60
Db 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSLFLALFSCULHPAASLEWNTSG 199
Qy 61 LYLVTNDCNSSIYVEADDDVILHTPGC 87

```

```

Db 200 LYLVTNDCNSSIYVEADDDVILHTPGC 226

RESULT 8
Q68870 PRELIMINARY; PRT: 3021 AA.
ID Q68870;
AC Q68870;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Gense for core, envelope and NS1 proteins.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RC Tissue=Serum;
RA Seelig R., Weber P., Seeling H.P., Ledger N., Bottner C., Renz M.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBSJ databases.
DR EMBL; X76918; CAA54244.1; -.
DR PIR; P01307; P01307.
DR PIR; P00401; P00401.
DR PIR; P00804; P00804.
DR PIR; S41288; S41288.
DR HSSP; P27958; 1ALV.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT CHAIN 736 1012 NS2 protein.
FT CHAIN 1013 1663 NS3 protein.
FT CHAIN 1664 1717 NS4a protein.
FT CHAIN 1 191 core protein.
FT CHAIN 1718 1978 NS4b protein.

```



```
FT CHAIN 1979 2430 NS5a protein.
FT CHAIN 2431 3021 NS5b protein.
FT CHAIN 192 383 envelope protein.
FT CHAIN 384 735 NS1 protein.
SQ SEQUENCE 3021 AA; 329093 MW; BF2B499AA55A58CF CRC64;

Query Match 48.3%; Score 87; DB 2; Length 3021;
Best Local Similarity 100.0%; Pred. No. 1.1e-77;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LVLTNDGSSSIVYEADDDVILHTPGC 87
DB 200 LVLTNDGSSSIVYEADDDVILHTPGC 226

RESULT 9
Q81258 ID Q81258 PRELIMINARY; PRT; 3021 AA.
AC Q81258;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RA Sakamoto M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; D17763; BAA04609.1; -
DR PIR; PQ0401; PQ0401.
DR PIR; PQ0804; PQ0804.
DR HSSP; P27958; 1HS1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.

PFam; PF01542; HCV core; 1.
PFam; PF01539; HCV env; 1.
PFam; PF01560; HCV NS1; 1.
PFam; PF01538; HCV NS2; 1.
PFam; PF02907; HCV NS3; 1.
PFam; PF01006; HCV NS4a; 1.
PFam; PF01001; HCV NS4b; 1.
PFam; PF01506; HCV NS5a; 1.
PFam; PF00398; Viral_RdRp; 1.
DR SMART; SMO0487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS00435; PEROXIDASE 1; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT CHAIN 735 1012 NS2.
FT CHAIN 1013 1683 NS3.
FT CHAIN 1664 1717 NS4a.
FT CHAIN 1 191 C.
FT CHAIN 1718 1978 NS4b.
FT CHAIN 1979 2430 NS5a.
FT CHAIN 2431 3021 NS5b.
FT CHAIN 192 383 E1.
FT CHAIN 384 735 E2/NS1.
SQ SEQUENCE 3021 AA; 329575 MW; 38712CCBC0C19562 CRC64;

Query Match 48.3%; Score 87; DB 2; Length 3021;
Best Local Similarity 100.0%; Pred. No. 1.1e-77;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 60
DB 140 VGAPVGGVARALAHGVRALEDGINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSG 199

QY 61 LVLTNDGSSSIVYEADDDVILHTPGC 87
DB 200 LVLTNDGSSSIVYEADDDVILHTPGC 226

RESULT 10
Q68728 ID Q68728 PRELIMINARY; PRT; 149 AA.
AC Q68728;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Precursor protein (Fragment).
OS Hepatitis C virus type 3a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 3.
OX NCBI_TaxID=31652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95052487; PubMed=7525693;
RA van Doorn L.J.; Kleter B.; Stuyver L.; Maertens G.; Brouwer H.;
RA Schalm S.; Heijink R.; Quint W.;
RT "Analysis of hepatitis C virus genotypes by a line probe assay and
correlation with antibody profiles.";
RL J. Hepatol. 21:122-129 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97201609; PubMed=9049395;
RA van Doorn L.J.; Kleter G.E.; Stuyver L.; Maertens G.; Brouwer J.T.;
RA Schalm S.W.; Heijink R.A.; Quint W.G.V.;
RT "Sequence analysis of hepatitis C virus genotypes 1 to 5 reveals
multiple novel subtypes in the Benelux countries.";
RL J. Gen. Virol. 76:1871-1876 (1995).
DR EMBL; L39298; AAA67821.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
```

DR Pfam; PF01542; HCV core; 1.  
 DR Pfam; PF01539; HCV env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT CHAIN <1 32 core protein.  
 FT CHAIN 33 >149 ei protein.  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 15872 MW; 7D48172927B98021 CRC64;

Query Match 47.2%; Score 85; DB 2; Length 149;  
 Best Local Similarity 100.0%; Pred.No. 9.2e-77;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LTDCNSNIIYVYADDVILHTPGICPCVQDGTSTCTWPTPTVAVKYVGATTASIRSHV 123  
 Db 44 LTDCNSNIIYVYADDVILHTPGICPCVQDGTSTCTWPTPTVAVKYVGATTASIRSHV 103

QY 124 DLLVGAATMCSALYVGDGCGAVFLV 148  
 Db 104 DLLVGAATMCSALYVGDGCGAVFLV 128

## RESULT 11

Q68176 PRELIMINARY; PRT; 128 AA.

DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Envelope protein (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3a;  
 RX MEDLINE=95146953; PubMed=7844535;  
 RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;  
 RT "Prevalence of hepatitis C virus sequence variants in South-East  
 RT Asia.";  
 RL J. Gen. Virol. 76:211-215(1995).  
 DR EMBL; U14212; AAC53901.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002519; HCV env; I.  
 DR Pfam; PF01539; HCV env; I.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 128 128  
 SQ SEQUENCE 128 AA; 13786 MW; FAA94C46DF80D0DA CRC64;

Query Match 43.3%; Score 78; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred.No. 8.7e-70;  
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VPTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRHQT 162  
 Db 51 VPTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRHQT 110

QY 163 QTCNCSLYPGHLSGHRMA 180  
 Db 111 QTCNCSLYPGHLSGHRMA 128

## RESULT 12

Q68178 PRELIMINARY; PRT; 128 AA.

ID Q68178  
 AC Q68178;

DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Envelope protein (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3a;  
 RX MEDLINE=95146953; PubMed=7844535;  
 RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;  
 RT "Prevalence of hepatitis C virus sequence variants in South-East  
 RT Asia.";  
 RL J. Gen. Virol. 76:211-215(1995).  
 DR EMBL; U14212; AAC53901.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002519; HCV env; I.  
 DR Pfam; PF01539; HCV env; I.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 128 128  
 SQ SEQUENCE 128 AA; 13835 MW; 04D7C339AB2AEC5F CRC64;  
 Query Match 43.3%; Score 78; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred.No. 8.7e-70;  
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VPTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRHQT 162  
 Db 51 VPTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRHQT 110

QY 163 QTCNCSLYPGHLSGHRMA 180  
 Db 111 QTCNCSLYPGHLSGHRMA 128

RESULT 13  
 Q80NE2 PRELIMINARY; PRT; 119 AA.  
 ID Q80NE2  
 AC Q80NE2  
 DT 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Envelope protein E1 (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22447295; PubMed=12560577;  
 RA Cantaloube J.F., Biagini P., Attoui H., Gallian P., de Micco P.,  
 RA de Lamballerie X;  
 RT "Evolution of hepatitis C virus in blood donors and their respective  
 RT recipients.";  
 RL J. Gen. Virol. 84:441-446(2003).  
 DR EMBL; AF515941; AA083248.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002519; HCV env; I.  
 DR Pfam; PF01539; HCV env; I.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 119 119

SQ SEQUENCE 119 AA; 13004 MW; D69FCA6EC118462F CRC64;  
Query Match 38.9%; Score 70; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 8.8e-62;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 111 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 170  
DB 21 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 80  
QY 171 PGLSGHRMA 180  
DB 81 PGLSGHRMA 90  
RESULT 14  
Q68185 PRELIMINARY; PRT; 128 AA.  
AC Q68185;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Envelope protein (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_taxid=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3a;  
RX MEDLINE=951146953; PubMed=7844535;  
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;  
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia";  
RL J. Gen. Virol. 76:211-215(1995).  
DR EMBL; U14221; AAC53910.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002519; HCV\_env.  
DR Pfam; PF01539; HCV\_env; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
FT NON\_TER 1 128  
SQ SEQUENCE 128 AA; 13859 MW; E896D3166D86184A CRC64;  
Query Match 38.9%; Score 70; DB 2; Length 128;  
Best Local Similarity 100.0%; Pred. No. 9.3e-62;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 111 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 170  
DB 59 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 118  
QY 171 PGLSGHRMA 180  
DB 119 PGLSGHRMA 128  
RESULT 15  
Q81277 PRELIMINARY; PRT; 180 AA.  
AC Q81277;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Core protein-envelope 1 protein (fragment).  
GN Name=core-E1;  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.

OX NCBI\_taxid=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93249436; PubMed=7683463;  
RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;  
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV type 3";  
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).  
DR EMBL; D14586; BAA03445.1; -.  
DR PIR; P01305; PC1305.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002519; HCV\_env.  
DR Pfam; PF01539; HCV\_env; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
FT NON\_TER 1 180  
SQ SEQUENCE 180 AA; 18887 MW; A24D1A0510CAA9F5 CRC64;  
Query Match 38.9%; Score 70; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.2e-61;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 111 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 170  
DB 111 YVGATTASIRSHVDLLVGAATMCSALYVGDGCGAVFLVGOAFTFRPRRHQTQTNCNSLY 170  
QY 171 PGLSGHRMA 180  
DB 171 PGLSGHRMA 180  
Search completed: December 1, 2004, 23:54.09  
Job time : 195 secs

***This Page Blank (uspto)***